An Assessment of the Biological Condition of the New Fork River, near the Pinedale Anticline Project Area: 2007

By:

Brett D. Marshall, M.Sc.

River Continuum Concepts

PO BOX 13

Willow Creek Montana 59760

! FINAL REPORT !

Prepared For Sublette County Conservation District, Pinedale, Wyoming

August 25, 2008

Version: 1.3

EXECUTIVE SUMMARY

This study compared sites along the New Fork River for negative influences of PAPA development on the ecology of benthic macroinvertebrates in 2007 and previous years. The data for this report were collected by the Sublette County Conservation District (SCCD) and included a modification of the field sampling method that allowed more informative analyses to be conducted on the data. The sites and methods used are based on the SCCD's baseline analysis of the New Fork River (Marshall 2005) and draws heavily upon the Wyoming Department of Environmental Quality (DEQ) standard procedures to assess the biological condition of streams using aquatic invertebrates (Hargett and Zumburge 2006).

In 2007 the sampling effort was increased to allow additional analytical methods to be used—and ultimately, after some calibration (next year) to replace a less-cost-effective sampling method. The results of using the new method, one that uses single Surber (SS) samples, indicated that it was an effective method with greater diagnostic power than the earlier methods (using composite Surber (CS) samples). However, a direct comparison of the methods showed there were some conflicting results between the methods. This underscores the need to continue with both methods next year, until the methods can be calibrated to compensate for the differences.

Site NF01, the farthest upstream site is influenced by a dam and appears to support invertebrates more typical of mountainous streams than the other sites. We recommended the replacement of this upstream reference because it represents an unattainable reference, makes the sites appear to be in worse condition than they are, and obscures some of real patterns in the distribution of species among the sites.

Four separate lines of analyses indicated that the site downstream from the pipeline (NF30) is differentially responding to sedimentation (as sand). The overall condition of the site is not significantly different from the "fully supportive" condition according to the newest edition WSII (Hargett and Zumburge 2006)—although scores were lower in the past, there were no significant differences in the WSII over time (CS methods). Thus the monitoring program has provided an early warning before the need of regulatory intervention. A reconnaissance of the river between NF30 and NF40 is recommended so that sources of erosion can be located and corrective actions, if any can be implemented.

- There was more sand in benthic samples from NF30 than from other sites (Fig. 3.2).
- NF30's sand was not inversely related to water speed (Fig. 3.3), suggesting that it is in a zone of active erosion.
- The biological metrics measured at this site indicated that sand-loving organisms dominated
- CCA indicated that sand correlated with the overall taxonomic structure at NF30.

1.0 Introduction

The purpose of this study is to characterize the biological condition of the New Fork River (Sublette County, Wyoming) and to assess impacts related to natural gas development. The study design is complex, but this complexity is required to discern any natural gas related impacts from natural variation in a very dynamic river system. There are several other forms of human influence (e.g., construction, sewer discharge) in the drainage as well as natural influences (e.g., stream size, substrata composition, mineral springs etc.) from which potential impacts will need to be differentiated. Thus, this study is more complex than many biological monitoring programs in the state, but it is that way because it needs to be to fulfill its purpose.

This year, we have introduced several new methods to the study, including new sites to help tease out the influence of the sediment laden East Fork River and an altogether new study reach to assess potential impacts of increasing development on the northern portion of the Mesa. Additionally, we introduced an improved sampling method to the study which needs to be evaluated relative to previous methods. Our previous studies of the New Fork River indicate that the initial impacts of gas development are *subtle*. The new methods should detect them sooner, and with some different analytical procedures, we can ensure the data are comparable to our earlier assessments. Once evaluated, these changes to the monitoring design will reduce both field and laboratory costs—while at the same time increasing the power of the study to detect changes in the river's condition. They will be discussed in greater detail in the methods section of this report. But for now the reader just needs to know that there will be new analytical methods in this report.

1.1 Biological Monitoring Background

Assessment of the biological condition of surface waters has become a key element in the comprehensive monitoring of water quality in the United States and beyond. States and federal agencies have been refining the techniques for regional assessment for about two decades, but the use of site specific designs (like the ones used here) began in the 1940's, and have been refined as computers have improved the kinds of statistical tests that can be applied to biological data.

Benthic macroinvertebrates are the most commonly used animal assemblage[1] used to describe ecological changes in rivers. "Benthic" is an adjective implying association with the bottom of streams or lakes. The "macro" part of the name means that, for much of the animals' life cycle, they are large enough to be seen without a microscope (though microscopes are required to identify them). "Invertebrates" are animals without backbones. Thus we are specifically monitoring aquatic insects,

[1] Assemblages are collections of species living together.

mussels, snails, worms, crayfish, crustaceans, mites, leeches and similar organisms. But the monitoring program does not use data from bacteria (they are micro-invertebrates) or fish (they are macro- vertebrates). These groups can also be used for biological monitoring, but their spatial temporal scales of response are not appropriate for the local scale of this project and would not allow impacts to be located.

Invertebrates are incredibly diverse and abundant. They are also critically important because they play critical roles in detrital food webs—including breaking down of complex organic material—and in transferring energy to higher trophic[2] levels by serving as food source. Together, these aspects make macroinvertebrate assemblages[3] excellent indicators of the overall health—or condition—of any ecosystem:

1. They are numerous enough to be effectively sampled.
2. They are diverse enough to exhibit response signatures.
3. They are important and relevant to all "higher" animals.
4. They respond rapidly enough to provide early warnings of problems.
5. Their response is to disturbance is recognized as important by many agencies.

For these reasons, benthic macroinvertebrates are often used to assess the effects of human activities to streams and rivers. Thus they may be used to describe the impacts of development and to describe the effectiveness of restoration (or mitigation). This is the rationale behind this study.

[2] "Trophic structure" refers to the level of organisms in the food chain (or food webs) and specifically refers to their roles in processing organic matter and moving its energy to other groups of animals. For example, algae, algae eating invertebrates, predatory invertebrates, and fish, might represent different trophic levels in a food web.

[3] Assemblages are collections of species living together.

1.2 PAPA[4] Biomonitoring Background

The Sublette County Conservation District initiated a baseline assessment of the New Fork and Green Rivers in Sublette County in years 2000 and 2001 respectively. The goals of the established monitoring network were to establish a baseline data set which might be useful to assess stream condition in the future. An early report (Marshall 2005a) found that there was significant within-site variation in assemblage structure using the Wyoming Department of Environmental Quality's (WY DEQ) bioassessment methods and that if the baseline assessment was going to adequately assess change annually, it needed to have replicate samples collected periodically. Without elevated monitoring effort it was likely that the monitoring program would fail to detect important changes for several years.

The first analyses we conducted relating to the PAPA examined water chemistry only at a few study sites (Marshall 2005b). This was a very cursory set of analyses designed to quickly get some information to the task group and it only used sites on the main-stem of the New Fork River (NF01, NF04, NF30, and NF19). It indicated that NF04 had elevated conductivity, which continued to increase downstream and that both TSS and turbidity were elevated at NF30. These results seemed to indicate that there was a PAPA-related sediment impact. However, when we completed a more rigorous analysis on all the sites (Marshall 2005a), we found that the site differences observed in TSS, Conductivity, and Turbidity were related to tributaries, not the PAPA. These results demonstrated the importance of quantifying natural changes in the river and its tributaries. That is, to ensure the monitoring program does not "cry wolf" by suggesting development is responsible for naturally occurring changes in the river's condition, or development elsewhere in the watershed, we need to monitor several sites in detail.

The current monitoring design has been developed to account for site differences that are not related to the PAPA, so that the subtle development-related changes can be teased out from other sources of variation. If necessary, this should give the task group time to make recommendations to mitigate impacts before the rivers condition is significantly degraded by the cumulative effects of minor impacts.

---

[4] Pinedale Anticline Project Area

1.3 Overview of this Report

This report is somewhat different from the previous reports monitoring the effects of PAPA development on the aquatic invertebrates of the New Fork River. The differences arise because we have more (and arguably better) data now for each year. Some of the analyses performed this year could not be performed previously because of limitations of the previous sampling methods[5].

We added several new sites to the study in 2007; some to ensure that the effects of development on the north-end of the Mesa are monitored, and some to help resolve natural community changes related to East Fork River (and other upstream influences) from potential effects of development on the Pinedale Anticline through Alkali or Sand Springs Draws or more directly near the development of the pipeline.

The field crew used slightly different methods to collect samples in 2007. To make the data comparable, we needed to add some additional statistical analyses. However, the scope of this report was not to directly compare methods of assessment thoroughly[6]. Still the use of different methods does complicate the forms of statistics we used and organization of this report. These methods are discussed in greater detail than will be of interest to most readers in the methods section of this report.

Different levels of sampling effort (samples and sites) in previous years prevent us from comparing differences among sites in both space and time simultaneously. Thus this report will perform the analyses separately. First we will examine differences observed among all the sites in 2007. The section will be divided into assessment of the overall invertebrate assemblage using multivariate statistical techniques. This will be followed by an assessment of 2007 differences in biological metrics and the influence of habitat on those metrics. Finally, we will compare the Wyoming Stream Invertebrate Index among sites.

---

[5] This has been discussed in previous reports and will be discussed in greater detail in the methods and discussion sections of this report.

[6] We will do this comparison in 2009's report when we have 2 years of data (2007, 2008) using both methods to ensure the findings were not unique—no sense conducting these comparisons several times, without adequate data.

2.0 Methods

2.1 Study sites

The sites examined for this study were located on the New Fork River to assess potential impacts from gas well development. The site on the East Fork River was sampled to account for the changes in community structure related to the addition of fine sediment and psammophilic[7] species to the New Fork River, from the East Fork. Many of the sites have been sampled as part of the Sublette County Conservation District's (SCCD's) on going effort to document baseline biological conditions of the county's surface waters (Marshall 2005a). Others were added over time to assess the specific monitoring needs related to development in the Pinedale Anticline Project Area (PAPA). This section of the report describes the location and characteristics of each site and is concluded with a brief discussion on how the sites together help provide a bigger picture of changes in the New Fork River drainage.

NF01 New Fork River

NF01 is located ~ 1½ miles downstream of the New Fork Lake Dam, near the outlet of the lake. The dam was breached in 1928 and was subsequently rebuilt. The New Fork Irrigation District maintains a gauging station on the dam. A Boy Scout Camp is located on the lower end of the lake approximately 1 mile from the dam. The upper portion of the watershed above the lake is a wilderness area in the Wind River Mountains. NF01 is the first site in the New Fork River watershed and serves as a control site for the Pinedale Anticline as there is no oil or gas exploration or development near or upstream of NF01. Baseline biological assessment indicated that this site had a unique biological and chemical composition from downstream sites (Marshall 2005a, 2005b). At this location, much of the river's flow is provided as through-fall from New Fork Lake Dam; the effects of dams on river food webs are well documented and numerous (nutrient sink, elevated seston, elevated filter feeders etc.). Additionally the confluence of Willow Creek and Duck Creek increase the conductivity of the New Fork River (Marshall 2005b) and alter the composition of the river's invertebrate communities.

NF04 New Fork River

NF04 is located south of Pinedale ~2 miles and is 50 feet downstream from the South Tyler Bridge. South Tyler is an access road for the PAPA. NF04, when established, was located upstream of the PAPA. A Wyoming Game and Fish Department fishing access and boat-launch are located at the sampling site. NF04 is also located downstream of the confluences of NF02 Willow Creek and NF03 Duck Creek; the confluence of these streams is believed to coincide with dramatic changes in the chemical and biological make up of the New Fork River (Marshall 2005a). Additionally, increased development on the north end of the Mesa may contribute potential runoff to the New Fork River upstream of this site.

---

[7] Sand-loving or sand-dwelling species are called psammophilic, or psammophilous.

NF17 East Fork River

NF17 is located on the East Fork River, ~0.125 miles upstream of the confluence with the New Fork River. The Wyoming Game and Fish Department Boulder Fish Rearing Station is located upstream of NF17. NF17 is located downstream of HWY 191 approximately 5 miles. The East Fork River at NF17 is a sand dominated system with active sediment transportation occurring continually. In combination with several other sites, this site serves as a reference to account for changes downstream because it is a natural source of fine sediments that change the nature of the New Fork River's substrate composition and biology.

NF19 New Fork River

NF19 is located on the New Fork River, upstream of the confluence with the Green River ~2 miles. The site is ~1½ miles downstream of a USGS gauging station and HWY 353. Badlands lie adjacent to the New Fork River upstream of NF19. NF19 is downstream from the PAPA. NF19 is the last sampling site in the New Fork River watershed. It serves as to describe the condition of the New Fork River before it mixes with the Green River, and to help characterize the nature of upstream changes. Thus, this site is the ultimate recovery zone site and we do not anticipate development in 2007 to reach this site.

NF30 New Fork River

NF30 is located downstream of most of the Anticline development and below several pipelines' hyporheic crossings. The site is located on BLM land and has been sampled since the year 2001. A gravel pit is located west of the sampling site. NF30 is located downstream of the confluence of the East Fork River (NF17) ~3 miles. Five replicated samples were collected at this site from 2004-2007. These samples represent the "study" community that was compared to NF40 and NF50 to describe the effects of development in the PAPA.

NF40 New Fork River

NF40 is located within the PAPA and above the pipelines' crossings. The site is below the confluence of the East Fork River (NF17), Sand Springs and Alkali Draws and upstream from NF30 by about 1.5 miles. Five replicated samples were collected at this site during the years 2004 to 2007, but it was not sampled prior to 2004. These replicated samples originally represented the "control" community for comparisons with NF30 to describe the effects of the Pinedale Anticline. The site is not an ideal control site because there is potential influence from Sand Springs Draw and Alkali Draw during runoff. This is likely to become more of a problem with planned (non-PAPA related) development in the upper reaches of Sand Springs Draw. Thus, this site is now considered a measurement of the combined influence of Sand Springs and Alkali Draws, when compared to NF50.

NF50 New Fork River

NF50 is located downstream of the confluence of the East Fork River (NF17) ~ ½ mile and upstream of Sand Springs and Alkali Draws. This site was established in 2007 account for the effects of the East Fork River on the biota of the New Fork River. This is important, because NF 40 may be influenced by elevated sediment expulsion from Sand Springs and Alkali Draws. If this were to occur, there would be no way to differentiate the effect from the influence of the sand-laden East Fork River. A Wyoming Game and Fish Department public fishing access and boat launch area is located at this sampling site. Only biological data is collected at NF50 based upon the decision of the Pinedale Anticline Water Task Group. No chemical data is collected at NF50.

NF60 New Fork River

NF60 is located upstream of the confluence of the East Fork River (NF17) with the New Fork River ~3/4 of a mile. NF60 was established in 2007 to describe the condition of the New Fork River before it is influenced by the East Fork River. This is important for documenting the influence of the East Fork River on the New Fork River at NF50. Only biological data is collected at NF60 based upon the decision of the Pinedale Anticline Water Task Group. No chemical data is collected at NF60.

NF70 New Fork River

NF70 is located upstream of the confluence of Pole Creek ~ ¼ mile and downstream of NF04 ~ 4 miles. NF70 was established to monitor any effects from exploration and development from the northern portion of the Pinedale Anticline Project Area. This site measures the cumulative changes related to the gas development and the influence of Pinedale's sewage treatment plant (Pine Creek) which may change over time if facility management should change. Only biological data is collected at NF70 based upon the decision of the Pinedale Anticline Water Task Group.

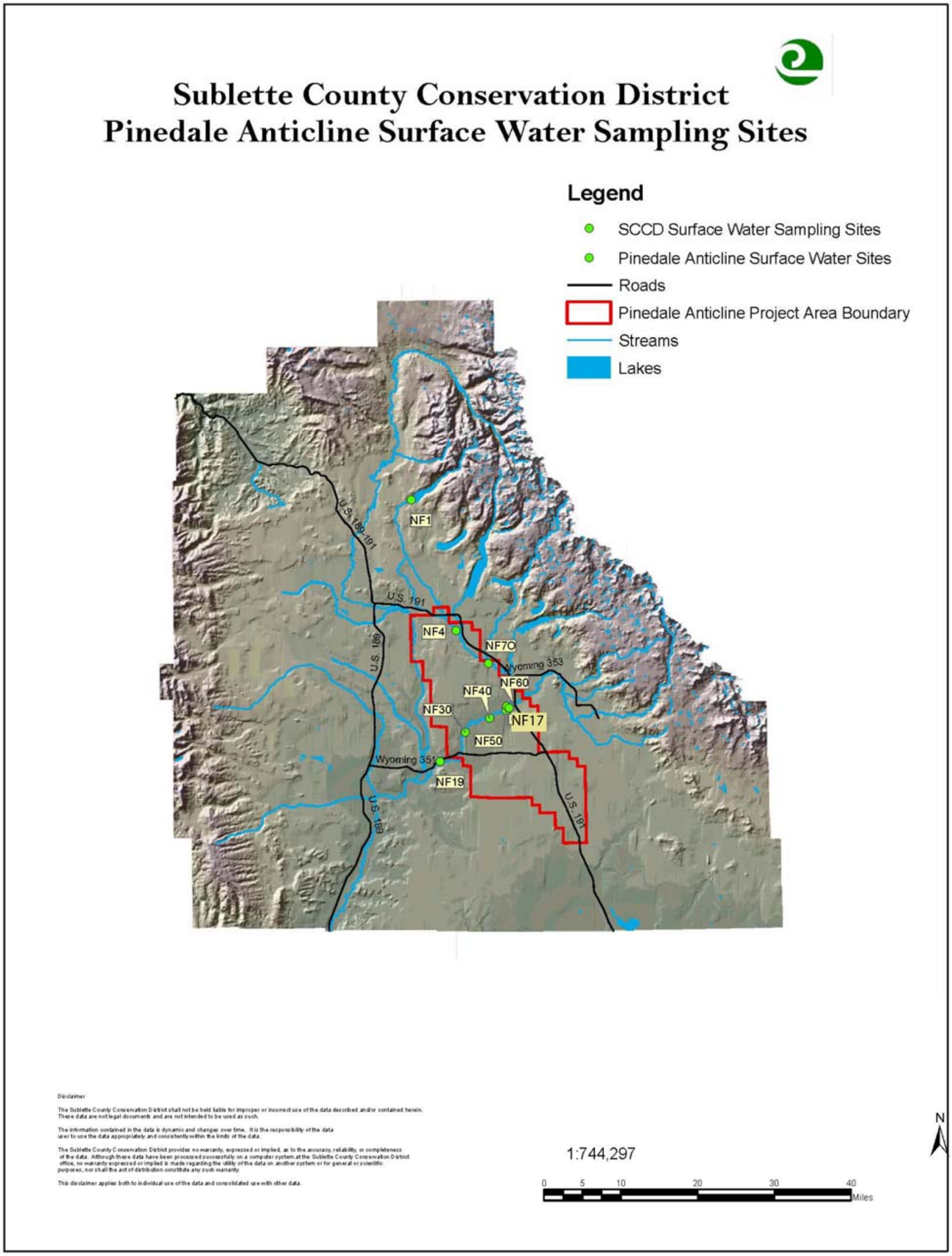


Figure 2.1. Map of Study Area and Sublette County, Wyoming. The location of all study sites are noted by green dots. The red polygon denotes the Pinedale Anticline Project Area (PAPA). This map was provided by the Sublette County Conservation District.

2.2. How the sites fit together

The study sites represent a cumulative gradient of effects. Thus, we have had to evolve this study from direct comparisons of two sites, to the use of some complex statistical procedures that account for natural gradients. We have identified several modes whereby the integrity of the New Fork River could be affected by run off from development on the PAPA. Currently there are two regions where potential effects of PAPA development are likely to accumulate as measureable impacts. We have separated these two areas into the Upper and Lower Study Areas to facilitate discussion. Graphs throughout the results section of this report have been bisected to clearly show the two study areas as well as the relative location of study sites along the downstream gradient. We have prioritized these locations based on the likely movement of surface waters during rain and snow melt events. This makes sense because these events are the most likely source of disturbance for surface waters—which are most likely to be in the form of eroded soil and sedimentation in streams. Additionally, if leachate or other industrial chemicals are spilt on soil, their eventual arrival in river systems is likely to correspond to runoff events. Note that there are no direct disposal effluents at this time.

The Lower Study Area

The primary concern was run off directly from development on the southeast section of the PAPA. This was the rationale for the early addition of NF30, NF40 was added latter to serve as a benchmark by which to gauge changes at this site. It soon became clear that this was not sufficient because we needed to account for changes from the New Fork River as well as potential impacts from runoff through Sand Springs and Alkali Draws, which enter the river downstream from the East Fork River and upstream from NF40. Thus several sites were added to account for this gradient.

NF17, on the East Fork River had traditionally been represented by a single bioassessment sample, which did not allow us to account for variation in the New Fork River that may be related to inputs from this sandy system. This site was recently augmented with replicate samples to allow us to include it as a spatial temporal variable in the statistical models. Conditions at NF 50 should result from a combination of the conditions at NF60 and NF17. The difference between NF50 and NF40 may account for runoff flushed through the draws. Direct runoff (as opposed to indirect runoff[8]) from the Anticline would be represented from changes in the condition of NF40 to NF30 (Fig. 2.2).

[8] Direct runoff is runoff going directly from development sites to perennial surface waters, indirect runoff occurs when industrial runoff enters surface waters indirectly—this happens when it flows to intermittent water bodies, such as Sand Springs or Alkali Draws, or enters surface water by way of hyporheic interaction with ground water (see footnote 9).

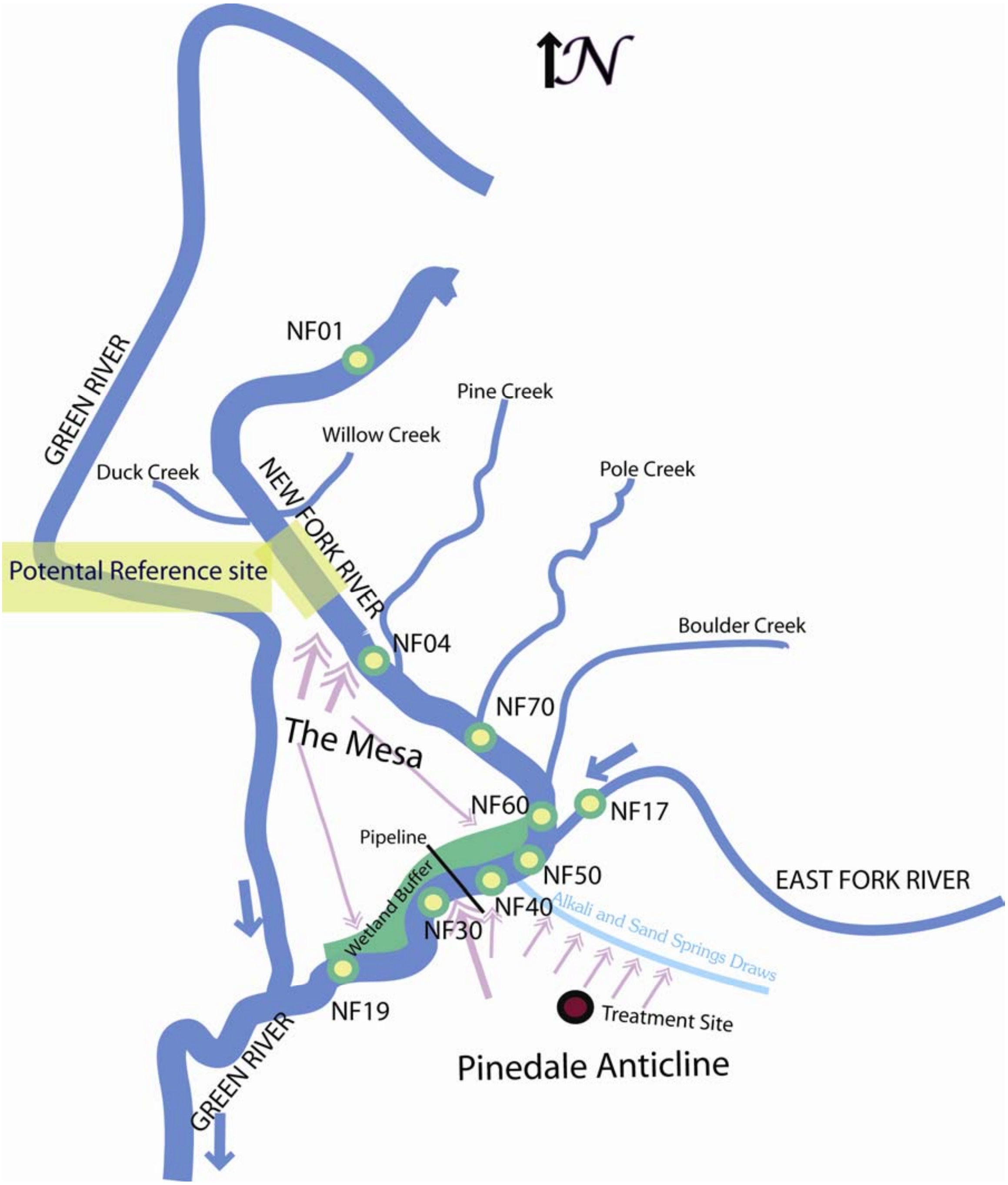


Figure 2.2. Study Site Schematic. This diagram shows the interrelationship among the location of study sites and potential sources of runoff in the PAPA. Double headed arrows indicate potential vectors of influence on the New Fork River from development on the PAPA. Although most of the run off on the mesa drains to the southeast, this run off encounters several wetland systems and is unlikely to actually reach the river. The sample sites are marked as circles on the river. The taxonomic composition of NF01 is significantly different from other sites because it is smaller and influenced by New Fork Dam. We have proposed replacing it with a reference downstream of the confluence of Duck/Willow Creeks, unless PAPA development is likely to begin affecting either of these tributaries as well. The ideal reference for the upper study unit would be in the highlighted area. However, if future development will impact Duck Creek, references should include a location above and below the confluence as well as at least one site on Duck Creek itself. We know from baseline analyses that these tributaries correspond with natural fundamental changes in the chemistry and biology of the New Fork River and to account for this will take a minimum of three study sites.

Most of the land comprising the Mesa drains to the southeast. Thus, it appeared likely that potential runoff and erosion could enter the river from the south-eastern edge of the Mesa. However, field investigations in 2006 indicated an extensive wetland system which will buffer the river from the effects of run off from the southeast edge of the Mesa (Fig 2.2). Thus, the most likely source of impacts to the lower study area is runoff from the southeast portion of the PAPA —directly (i.e., pipelines or site-runoff to the northwest) or indirectly (via the draws). Unless, that is, development on the northwest side of the river encroaches into the riparian zone—in which case sediment impacts are likely due to direct runoff, hyporheic[9] disturbance, and reducing the capacity of the wetlands to buffer upland impacts.

The Upper Study Area

Although most of the runoff from the Mesa flows southeast, there is an area on the northern edge of the Mesa which drains northerly. This area was unaccounted for by previous monitoring efforts and 2007 was the first time the PAPA monitoring project focused on potential impacts in this area. In 2007, three sites (NF01, NF04, and NF70) were sampled to account for changes in this study area. Although the area is smaller than the lower study area, the gradients are as complex as those occurring in the lower project area. NF01 was selected as a reference site, although it is influenced by an impoundment and is smaller than much of the river. NF04 is different from NF01, because of natural influences from upstream tributaries (Willow Cr. and Duck Cr.). The results of this report indicated that NF01 should be replaced as a reference. Therefore NF04 was used as an upstream reference to compare to NF70. But NF04 could also be influenced by upstream influences from the north end of the PAPA. This is not ideal but can be improved with another reference site.

NF70 integrates the effects of several smaller drainage systems off the Mesa, but is also influenced by Pine Creek and Pole Creek. Both of which were found to convey some anthropogenic influence to the New Fork River. This makes it difficult to assess PAPA related influences from other anthropogenic stressors, and may require replicate sampling on those tributaries as well (as mentioned in last year's report). Meanwhile, the full benefits of this site will not be fully realized until next year's report, when we have a second year of data to compare to data collected in 2007. Over the long-term, the temporal changes occurring at this site relative to NF04 and NF60 (the upstream, site of the lower study unit) will be important to diagnose changes within the upper study area and the lower study area as well.

---

[9] The hyporheic zone is the subsurface flowing portion of the river and an important zone of interaction between surface water and ground water. Many immature insects rear early life stages in hyporheos. Abandoned stream channels often have seasonal or perennial hyporheic flows and can carry stressors back to surface waters.

2.3. Field methods

The first part of this report details the differences occurring among sites in 2007. This is different from the previous year's analysis which examined the differences among sites over time. We also examined changes over time in this report, but this required some statistical standardization of the 2007 data for comparability and will be discussed later. The deviation from previous years' analysis was necessary to allow us to use more powerful[10] statistical methods than we could previously. The only way we could replicate and use the SCCD baseline, which was based on Wyoming DEQ field methods, was to replicate the collection of "composite samples" (CS). While this offered several advantages (regional reference criteria, and reference to SCCD's Baseline data set), it also grossly limited types of analyses that could be done. Moreover, it limited our ability to tease out natural sources of variation (such as water velocity) which can confound an assessment's ability to infer cause-and-effect. Therefore we introduced non-composite samples, called "single samples" (SS), to the monitoring program this year. Our ultimate goal is standardize these results so that (hopefully) we will be able to replace the CS with SS (though this maybe several years off). The specifics of the methods are discussed below.

Single sample methods

The new sampling method for 2007 included eight single Surber samplers from each site, each of which was processed individually in the laboratory to Wyoming Department of Environmental Quality's Standard procedures (e.g., Stribling et al. 2000). The procedures deviated from DEQ's Standard methods, which were in earlier PAPA assessments, in that DEQ usually "composites" all eight samples into a single sample representing the site. In order for the SCCD field crews to collect replicates[11], they had to actually disturb 40 ft^2 and remove all insects and debris from the bottom of the river. By keeping the samples separate, we can correlate them with environmental variables and increase the statistical power of assessment. Furthermore, we can (and did) electronically composite the data to make samples comparable to WY DEQ standards and our earlier PAPA dataset.

Single samples were collected using a stratified random sampling regime where near-substrate flow measures were used to ensure that the samples from each site fell within a uniform range of flows. This procedure is important for several reasons. First, it ensures that flows are uniform among sites. We know that near substrata flows can

---

[10] "Power" is actually a technical term in statistics. More powerful designs have a low probability of type-2 statistical error. In the case of this report, this essentially translates to a design that is more likely to properly diagnose impacts as they actually occur.

[11] Replicate samples are required to statistically compare sites to determine if they are different from each other.

account for a very large amount of variation in aquatic invertebrate assemblages (e.g., Hart and Fonseca 1998, Hart and Finelli 1999) and we know that gas development is not likely to alter flow regimes. So by sampling consistent velocities we prevent this from producing confounding results.

In addition to ensuring flow consistency among velocities, sampling a range of flows at sites allows us to account for the effects of velocity on biological measures statistically. For this to succeed, each site needs to have a sufficient range of velocities to encompass a meaningful amount of biological variation, and we need the range to be similar among all sites. This technique is called Analysis of Covariance (ANCOVA; Zar 1999) and it uses the General Linear Models (GLM) algorithm (Wilkinson 2006) common among statistical software packages. We used SYSTAT v.11 statistical Analysis software for most analyses.

This technique also allowed us to relate other habitat variables directly to biological measurements. Many of these could be related to natural gas development—or due to natural variation. For example, for each sample, the field crew measured the relative substrata size distribution, and embeddedness. Thus, some statistical tests were run comparing sites without accounting for these variables, and then ran them again so that the influence of fine sediments on biological differences among sites can be resolved. These methods also allowed us to use several multivariate statistical procedures—which will be briefly discussed under statistical methods.

These samples are called "single samples" in graphs and tables to differentiate them from composited samples. Readers should not confuse this reference to un-replicated samples, because wherever "single samples" were collected, we collected eight stratified random statistical replicate samples. Thus, the values reported from "single samples" are means of eight individual samples and their associated 95% confidence interval.

Composite sample methods

SCCD also collected some composite samples similar to those collected before. This was especially important because the focus of this report is to detect PAPA related change and to complete the study we need data that we know are comparable to the baseline data collected from the SCCD baseline study and previous years of PAPA monitoring. These data sources were also used in this report. Each composite sample is made of eight randomly-selected (scattered all over the river bottom) single, square-foot benthic samples that were poured together in the field, preserved and sent to the laboratory for analysis. The problem with this method is that the sample can never be re-separated to correspond to other variables (such as flow, embeddedness, depth, and particle size distribution). Another disadvantage is that this is an incredibly ineffective method to replicate; to attain only five replicates requires the collection of 40 samples.

The advantage of this method is that DEQ's standard assessment tools were calibrated using this method. By using this method we can use the Wyoming Stream

Invertebrate Index and have the condition ratings actually rate the sites relative to regional reference criteria.

Composite samples were only collected from sites with a history of being sampled with this technique. As with earlier assessments, SCCD collected five replicates using this method so that we could compare from site to site and year to year.

2.4 Laboratory

Biological metrics data from 2000-2004 were entered and validated by SCCD personnel and sent to Brett Marshall for analysis by a professional stream ecologist. Certified professional laboratories completed all the laboratory analyses and trained SCCD staff collected all field measurements. Thus, this report meets the requirements for credible data defined by the State of Wyoming. Most biological data from 2004-2007 were generated from raw taxonomic data by EcoAnalysts, Inc. according to the taxonomic standards set forth by WY DEQ.

Single samples (as discussed above), which are not used by WY DEQ at this time, were subsampled to allow the identification of 200 organisms to the genus species level[12], including midges and worms. If specimen condition or maturity prevented this level of taxonomy they were identified to the lowest practical taxonomic level. If the single samples contained fewer than 200 individuals, the entire sample was identified.

Composite samples were subsampled to ensure that 500 individual organisms were identified to the specified levels defined by WY DEQ. This includes species level identification for all taxa, including midges and worms when specimen maturity and condition permits. If the composite samples contained fewer than 500 organisms, the entire sample was sorted.

A complete quality assurance report was submitted to SCCD with the data and indicated that, similar to previous years, all invertebrate laboratory procedures met or surpassed the sorting and taxonomy standards required for these types of data.

[12] This is the highest level of taxonomic resolution possible for macroinvertebrates.

2.5 Analysis

Habitat variables

The area contained within each SS benthic sample was described to provide sample-specific habitat data. These data were collected and recorded by SCCD during field collection and added to the analytical data set. These measures included depth, flow (6/10 depth and near substrata), % size composition of inorganic substrata (Wentworth 1922), and embeddedness.

In similar assessments (Marshall 1997, 1998, 1999, 2007a) conducted for the Academy of Natural Sciences, I have found that compiling substrate size distribution data in to a particle size index has certain advantages. It correlates well with biological metrics and avoids problems with autocorrelation caused by using all the measures (which are proportional to each other). The index I have used in the past weighs the percentage of each substrate size class, relative to the suitability for invertebrate colonization. For example, many invertebrate species do not like sand—it moves in the river flow and could bury them or grind them up. Similarly fine gravel frequently moves (less than sand, but more than larger particles. Optimal balance of providing surface area and stability is attained by the cobble-sized particles. Boulders are stable, but have less surface area per unit volume to accommodate diverse communities.

IPI = 0*fines + 1*Fine Gravel + 2*Coarse Gravel + 3*Pebble +4*Cobble + 1*Boulder

Biological variables

Biological metrics are values calculated from the taxonomic data set (which is a list of the species collected and their abundance) because they summarize the changes in species composition in terms of changes in ecological function. Metrics were used as the response variables for most analyses. This was necessary because the abundances of species change naturally though time and in space due to changes in the environment, inter-species competition, and other factors. Ecological theory predicts that the functions performed by these species should be conserved—unless the ecosystem's function in impaired. That is, the abundance of each species may change naturally as a response to climatic variation or natural biological cycles, but usually a reduction in the abundance of one species is accompanied by an increase in the abundance of similar species. Thus, measures like the relative abundance of collector-gatherers should be more consistent than the abundance of individual species comprising the collector-gatherer guild. This is how metrics reduce the variability in species abundances by summarizing functional changes. The metrics compared in this report are discussed briefly below. The WSII mentioned throughout the text is the Wyoming Stream Invertebrate Index developed by Wyoming DEQ. It is a method of combining metrics to summarize the ecological condition of streams and rivers in the state and will be discussed in the next section (Section 1.5).

Taxa Richness is a very common metric that is used to describe the function of terrestrial and aquatic ecosystems. The measure is calculated by counting the number of different species (or similar kinds) in the sample. For aquatic ecology, the underlying philosophy is that more species can live in clean water than in polluted water. Therefore, higher values of Taxa Richness indicate a "healthier" condition and lower richness values may indicate an impaired condition.

The orders Ephemeroptera, Plecoptera, and Trichoptera (mayflies, stoneflies, and caddisflies, respectively - EPT) are generally considered to be more sensitive to disturbance than other organisms. Although not universally true, many of these organisms need cool, flowing water with high oxygen and low ion concentrations year-round. Thus, one of the most popular metrics in the United States today is the EPT index, which is the taxa richness of these three sensitive orders (e.g., Lenat and Penrose 1993). Because these orders do not always respond uniformly, many states—including Wyoming—have started using the richness of each of the EPT orders as separate metrics. Thus, three of the metrics used in the WSII are the richness of Ephemeroptera, Plecoptera, and Trichoptera represented separately. We used combined EPT richness metrics to compare NF30 directly to NF40, but the richness of the individual EPT orders was used to calculate the WSII and compare samples to the regional reference condition.

The metric "percent abundance of Trichoptera" is based on the same philosophy as the other EPT measures. Many states, including Wyoming, exclude the family Hydropsychidae from this calculation because members of this family are generally more tolerant than most caddisflies. Thus, this metric excludes these taxa for this study and is calculated by dividing the total number of non-Hydropsychidae Trichoptera by the total number of invertebrates identified and multiplying by 100.

In the "new" WSII (Hargett and ZumBerge 2006), called WSII-2 hereafter, the metric, "%Trichoptera excluding hydropsychids" (above), was replaced by the metric "%Non-Hydropsychidae of Trichoptera. The calculation for this metric is to divide the abundance of Non-hydropsychid caddisflies by the total abundance of caddisflies (not the total abundance of all invertebrates) and multiply by 100. The philosophy behind this metric is similar to the %Trichoptera-excluding Hydropsychidae metric.

The abundance of chironomid midges is often used as an indicator of environmental perturbation because there are 4000 species known from the northern hemisphere. Some of the common species are very tolerant to certain stressors and reach very high abundances when densities of predatory insects or competitors are reduced in polluted waters. This metric responds to organic enrichment and sedimentation, as well as acid mine waters. Specific taxa comprising the chironomid assemblage can be particularly useful for describing the causes of changes in multi-metric indices (like the WSII) and other metrics.

North American streams are normally dominated in abundance, richness, biomass, and production by aquatic insects. The notable exceptions are high-mineral springs and highly disturbed streams. Thus, high numbers of non-insect invertebrates often indicate that streams are stressed, or that there are unusual circumstances governing the community structure. Some non-insects, such as the ubiquitous amphipod *Hyallela* sp., are very tolerant of stress from high temperatures and elevated salinity. Others, like aquatic earthworms are tolerant to organic or inorganic sedimentation. Thus, specific taxa can be useful to help diagnose the causes or nature of anthropogenic perturbations.

The Biological Community Index, Community Tolerance Quotient (BCI CTQ) is an index that is calculated by combining the tolerance of invertebrate species to ions with the expectations for the region. High values indicate a community of high tolerance, whereas low values indicate a community dominated by sensitive organisms. High values indicate a community dominated by organisms preferring higher conductivity waters.

In the WSII-2 (Hargett and ZumBerge 2006), BCI is replaced by the Hilsenhoff Biotic Index. This index uses the weighed abundance of organism's tolerance to organic pollution to score the sample from zero to ten. Low scores mean that most organisms are very sensitive to pollution. High scores indicate that most organisms in the sample are very tolerant to pollution. For example, the rat-tailed maggot, Syrphidae, has a tolerance of 10, and can be collected from sewage treatment lagoons. Thus, high values, usually indicate a polluted condition, where as low values usually occur in clean, cool, low-sediment waters. It is important to remember that this metric has been specifically developed to describe the effects of organic pollution. It also responds to some forms of sedimentation, but is pretty ineffective to describe changes in pH, metals, or other ionic changes in water quality.

Semivoltine taxa are taxa that require more than 1 year to reach reproductive maturity. These taxa have generation times of 2 or more years. We found that some of the earlier data (Bollman 2004) confused long-lived taxa with semivoltine taxa and therefore, I recalculated the metric for this analysis. The semivoltine taxa in these samples included some stoneflies, and dragonflies in the family Gomphidae. Earlier calculations had included some long-lived beetles. Although these taxa are long-lived, they may reach maturity in 1 year and should have been classified as univoltine. The rationale for this metric is that these taxa take a long time to reproduce and re-colonize habitats. Thus, streams that are subject to frequent severe, but intermittent, disturbances will support very few semivoltine taxa. The WSII-2 (Hargett and ZumBerge 2006) specifically excludes beetles from the calculation of Semivoltine richness—eliminating potential confusion by neophytes between the terms, "semivoltine" and "long-lived" (some beetles are actually semivoltine).

Wyoming Stream Invertebrate Index

The Wyoming Stream Invertebrate Index has recently been revised from the criteria used in earlier reports and baseline development (Stribling et al. 2000, Jessup et al. 2002, ZumBerge 2004 Pers. Comm.), to a new form (Hargett and ZumBerge 2006). However at the time of this report we only have limited data available and cannot re-describe the entire baseline conditions of the New Fork River. Therefore, we use both the WSII-1 (Stribling et al. 2000, Jessup et al. 2002, ZumBerge 2004 Pers. Comm.) and the WSII-2 (Hargett and ZumBurge 2006) in this report. The WSII-1 is used to compare with baseline conditions described earlier (Marshall 2005a), whereas the WSII-2 is used to describe the current results in the context of current Wyoming Monitoring Standards. At this time we are unsure of which of the two WSII's versions most accurately reflects the condition of sites in Sublette County, but the WSII-2 is based on much more data. This monitoring project's goal is specifically to detect changes related to PAPA development; this is a decidedly different goal than comparing samples to regional reference conditions (the goal of the WSIIs). We used the WSIIs because they have been found to respond to human stressors as have the metrics of which they are composed. Thus readers should understand that for this report, the focus is on describing change in the WSIIs and their metrics, rather than the actual condition score derived by the WSIIs. However, when the results of the WSIIs aid the interpretation of change, they are referred to in the text. The sites used in this report are all located in the Wyoming Basin Ecoregion (Fig. 1.3) and used reference criteria appropriate for that region.

The Old WSII scores 10 metrics (Table 1.1) between zero and 100, then averages the scores to arrive at a numerical value (0-100) describing the ecological condition of the river. WSII values near 100 describe streams that are very similar to the regional reference streams upon which the WSII is based. Values near zero indicate streams that deviate significantly from reference conditions. The numerical scores are also used to derive narrative condition classifications (Table 1.2) to infer the "health" of the stream communities—again, compared to the regional reference. For example, scores close to zero mean that the stream appears severely impaired relative to the reference criteria collected in the year 2000-2001—and the site would be classified as "Very Poor" condition.

.

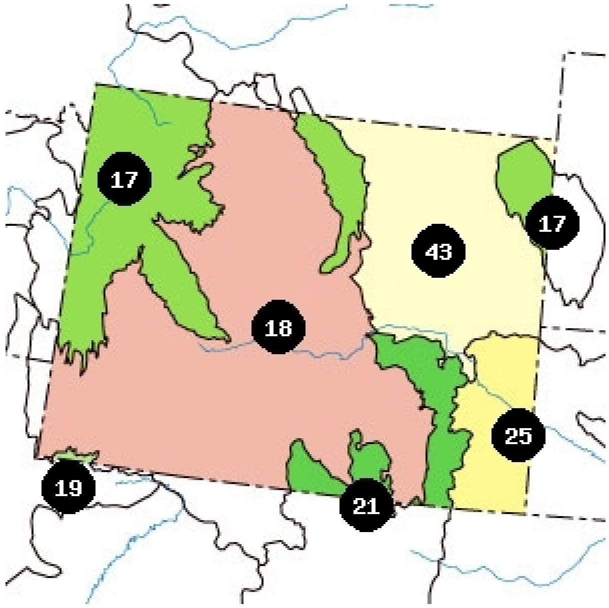


Figure 2.3. Wyoming Ecoregions. All the sites sampled in the New Fork Basin were in the Wyoming Basin ecoregion—denoted by the number 18. All streams in region 18 use the same criteria for the Wyoming Stream Invertebrate Index (WSII-1 and WSII-2).

Table 2.1. WSII-1 Scoring Criteria for the Wyoming Basin Ecoregion. Metric scores are inserted in place of the word "metric" in the scoring formula column. Values in the 5th or 95th percentile are based on the earlier survey of reference streams through out the Wyoming Basin Ecoregion. The average score for all 10 metrics is then used to attain the mean quality relative to the Wyoming Basin streams comprising the reference.

Wyoming Basin Ecoregion Rating Formulae (OLD WSII[13])

Metric	Scoring formula	5th or 95th percentile
Total taxa	100*metric / 95th%ile	45
Ephemeroptera taxa	100*metric / 95th%ile	9
Plecoptera taxa	100*metric / 95th%ile	5
Trichoptera taxa	100*metric / 95th%ile	10
% Plecoptera	100*metric / 95th%ile	13
% Trichoptera (no Hydropsychidae)	100*metric / 95th%ile	31.3
% Non-insects	100*(55 - metric)/(55 - 5th%ile)	0.5
% Scrapers	100*metric / 95th%ile	31.8
BCI CTQa	100*(110 - metric)/(110 - 5th%ile)	62.6
Semi-voltine taxa	100*metric / 95th%ile	7

[13] Based on Stribling et al. 2000, Jessup et al. 2002, ZumBerge pers. comm.. 2004

Table 2.2. WSII-1 Scoring Narrative Condition Criteria for the Wyoming Basin. Once the metrics are scored according to the criteria in Table 1.1, the average score is used to determine the narrative condition classification from the table below. For example, using the WSII-1 a mean Metric score of 85 would be called "Very Good," whereas a mean metric score of 43 would be called "Fair."

Condition	Minimum	Maximum
Very Good	80.5	100
Good	60.9	< 80.5
Fair	40.6	< 60.9
Poor	20.3	< 40.6
Very Poor	0	< 20.3

The WSII-2 (Hargett and ZumBerge 2006) is very similar to the WSII-1 (Stribling et al. 2000, Jessup et al. 2002, ZumBerge 2004 Pers. Comm.) except that some of the metrics have been changed, scoring criteria adjusted, and nature of the narrative condition criteria was altered. The specific metrics are discussed in section 1.4 (above). The scoring criteria (Table 1.3) are similar—they rate score metrics between zero and 100 based upon comparison with regional reference criteria. The new condition classes are "Full-support", "Intermediate," and "Partial / Non-support" of aquatic life use. "Full-support" of aquatic life use indicates that the site (or sample) exceeds the 25th percentile of reference streams. Scores below the 25th percentile of aquatic life uses were split into three equal classifications. The "Intermediate" category is the upper 1/3 of reference sites below the 25th percentile. This roughly translates to the range of the 50th-25th percentile of reference streams. Ancillary data need to be provided to interpret the condition of these sites—although they still exceed the condition of about 50% of the reference sites. Sites that score below the 50th percentile of the reference streams are said to have suffered "substantial anthropogenic perturbations" near the sample location (Hargett and ZumBerge 2006, p.25).

Table 2.3. WSII-2 Scoring Criteria for the Wyoming Basin Ecoregion. Metric scores are inserted in place of the word "metric" in the scoring formula column. Values in the 5th or 95th percentile are based on the earlier survey of reference streams through out the Wyoming Basin Ecoregion. The average score for all 10 metrics is then used to attain the mean quality relative to the Wyoming Basin streams comprising the reference.

Wyoming Basin Ecoregion Rating Formulae (WSII-2[14])

Metric	*Scoring formula*	*5th or 95th percentile*
Ephemeroptera taxa	100*metric / 95th%ile	9
Trichoptera taxa	100*metric / 95th%ile	9
Plecoptera taxa	100*metric / 95th%ile	6
% non-insects	100*(64 - metric)/(64 - 5th%ile)	0.4
% Plecoptera	100*metric / 95th%ile	22.3
%Non-Hydropsychidae of Trichoptera	100*metric / 95th%ile	31.3
%Collector-gatherer	100*(96 - metric)/(96 - 5th%ile)	100
% Scrapers	100*metric / 95th%ile	38.6
HBI	100*(8.3 - metric)/(8.3 - 5th%ile)	1.9
Semi-voltine taxa (No Coleoptera)	100*metric / 95th%ile	5

Table 2.4. WSII-2 Scoring Narrative Condition Criteria for the Wyoming Basin. Once the metrics are scored according to the criteria in Table 1.3, the average score is used to determine the narrative condition classification from the table below. For example, using the WSII-2, a mean metric score of 56 would be called "Full-Support," whereas a mean metric score of 50 would be called "Intermediate." A mean metric score of 34.5 or less would be called "Partial-Support" or "Non-support." The meaning of these terms is discussed in the text.

Condition	Minimum	Maximum
Full-Support	>51.9	100
Intermediate	51.9	34.6
Partial / Non-Support	<34.6	0

[14] Hargett and ZumBerge 2006.

Statistical Analyses

The goal of this monitoring project has a different goal than comparing with other streams throughout Wyoming; we want to know if gas development in the Pinedale Anticline Project Area is changing the biology of the New Fork River. This is a much more complicated question than can be answered by the WSII's narrative condition criteria. We know from past experience that there are some natural deviations from the regional references of the WSII. However, our study design was developed to allow us to use the WSII to test sites NF30 and NF40 for changes related to the Anticline. Additionally, we use metrics that we have historically been useful in the New Fork River (Marshall 2005a).

As with previous years, the method we used to test the CS[15] Sample metrics (including the WSII(s)) for change is called two-way analysis of variance. (2-way ANOVA). This tests the data specifically for differences among different levels of "treatments" and interaction-effects (Zar 1999). In this study the treatments are SITE (NF30, NF40) and YEAR (2004, 2005, 2006, 2007). The ANOVA used the within site averages and variance to determine the likelihood that the levels of each treatment are sufficiently similar to be considered statistically representative of the same population of data. In application, a P-value (probability) that is small means that there is a low probability that the observations are sufficiently similar to belong to the same "group." The convention among research scientists is to use a critical P-value of $P=0.05$ (5%) as the decision threshold. Thus, if $P<0.05$, there is >95% likelihood that the compared groups are not homologous. Another way to say this is that the probability of "type-1 statistical error" is less than 5%; we have a < 5% chance to incorrectly conclude that homologous groups are not actually homologous.

Although a very low type-1 statistical error is paramount for sound science, it has been criticized for environmental monitoring because it may cause real and important environmental changes to be obscured by natural variation. To avoid this conundrum, we also examined all metrics with a more-liberal P-value ($P<0.10$) and called these changes "marginally statistically significant" or "marginally significant". When these terms arise they mean that the result was not significant at the 95%-level, but was at the 90%.

The 2-way ANOVA's results can indicate three kinds of differences: (1) significant difference between the two sites, (2) significant differences among the four years, and (3) significant interactions between treatments (YEAR and SITE). We know in advance that natural differences occur between Sites and among Years (e.g., Marshall 2006). Describing these is important. But for this study an impairment signature occurs when sites respond differently than references overtime. Thus, a statistically significant interaction ($P<0.05$) could be indicative of a sites impairment—or recovery—related to PAPA development (Zar 1999). There are a number of reasons that the sites do not

[15] CS=Composite sample, see field methods.

respond the same through time, and when I observed a significant interaction term, I examined the data more closely to determine the cause for differential response of the metric. These are discussed in the results.

New Analyses in the 2007 Report

This report used methods that allowed more efficient use of statistical analyses than in previous reports. Among these were methods that allowed us to include habitat measures in analyses of biological data. The first new[16] method is called "multiple regressions." It uses the General Linear Models (GLM) algorithm (available in most statistical software) and uses metrics as response variables and all the habitat variables as predictors. The modeling procedure then removes predictor variables (i.e., flow, particle size, embeddedness etc.), which, when tested, do not explain a significant amount of variation in a specific response variable (i.e., metric). These are removed one at a time until only the variables that significantly explain variation in the metrics tested. For this reason the procedure is called a "backwards step-wise multiple regression modeling algorithm," but throughout this report we call it simply the "multiple regression." This procedure was especially useful to describe which metrics appeared to correlate with water speed or sedimentation.

The second new analytical method used was analysis of covariance (ANCOVA), which also uses the GLM algorithm. In this case, we used it two ways. First, it tested significant differences in the response slopes of habitat variables responding strangely (velocity vs. % sand). The second use of ANCOVA was to statistically adjust the mean metric values to correct for the influence of water velocity—which is not related to PAPA development, to make PAPA influences more apparent. When this procedure significantly altered the results of the multiple regression or ANOVA, it was discussed in the results.

The third was a combination of community ordination techniques. Ordination methods plot the species abundance data in an N-dimensional hyper-space determined by the number of the species. Mathematically these techniques draw upon matrix algebra to tease trends from the variance-covariance matrices of the species abundance data. But you don't need to under stand this to appreciate what the methods produce. In the figures below, the red and blue spots are samples from different habitats and their abundance of Species -1 and Species-2 (Fig 2.5). A relatively non-technical review of community ordination is provided by Manley (2006); this text covers the basic matrix algebra and application of basic and advanced ordination techniques, including the Detrended correspondence analysis used in this report. However, it does not cover the more advanced Canonical Correspondence Analysis (CCA; ter Braak 1986, 1994,

[16] "New," in this case means, "not used in previous PAPA reports." It does not mean we used experimental methods; these techniques have been around for many years.

McCune and Mefford 1994), used to draw correlations between species abundance data and physical measures. Manley (2006) does provide a discussion of Canonical Correlation Analysis (cca)—which is similar to the CCA we used except that it (cca) assumes a linear response of species to physical measures, whereas CCA assumes a unimodal response to physical variables—which is more appropriate in this case.

FYI: to understand the basic idea behind ordination techniques, read this section… otherwise skip ahead.

To understand how these ordination techniques work in general, it is helpful to consider a simple example (Fig. 2.4). Consider samples collected from two locations containing different abundances of two species. The location from which samples were collected is indicated by the color of the point in the graphics. Traditional statistics would test the abundance of species 1 or species 2, or both. But if you project the data cloud to either axis, as traditional parametric statistical analyses would, you'll observe a significant overlap in the range of abundances of species-1 and species-2.; they probably would not be statistically significant even though a pattern appears to exist (Fig 2.4, first frame). Ordination techniques generate a new set of axes that explain the most variation in the data set (Fig. 2.4, second panel) from which comparisons can be made; these become new ordinal axes (Fig 2.4, third panel). In most ecological studies there are many more than 2 species to examine; the SCCD has about 200 species, but you cannot visualize a graph in 200-dimention-space. Therefore, ordinations plot the axes through multiple dimensions and builds mathematical expressions of multiple variables to serve as new axes. Thus, the axis that explains the greatest amount of variation in the samples becomes a mathematical expression of the abundance of all species and is called **Axis-1**. The axis that describes the second-most amount of variation among the samples becomes **Axis-2** and so on. These axes are unit less combinations of many variables (in this case variables are species) so they are just referred to as **Axis-1**, **Axis-2**, and **Axis-3,** which might seem unusual to some readers.

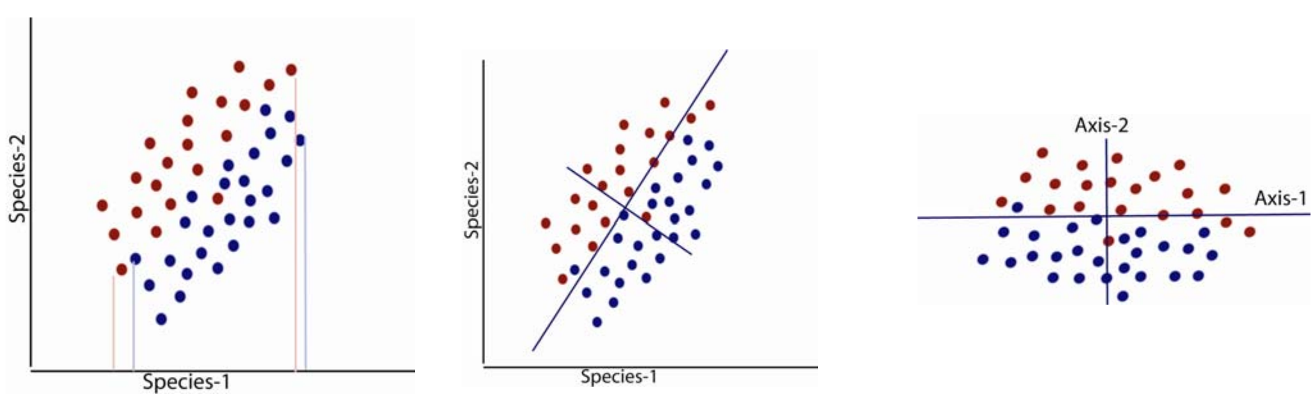


Figure 2.4. Ordination Example. The three panels illustrate a generalized ordination from left to right.

The actual methods used in this report were called Detrended Correspondence Analysis (DCA) and Canonical Correspondence Analysis (CCA). DCA uses only the number of species in each sample to develop groupings. DCA is very similar to the method described above, except that it uses non-linear (parabolic) to compensate for non-linear response of species to environmental gradients.

CCA uses the basic methods of DCA and constrains the ordination by the correlation with a habitat variable matrix. This allows an overlay of influential habitat variables as vectors of influence. If some sites collectively gather on the right side of the ordination plot, the vectors of influence allow investigators to develop a plausible environmental rationale for their shift in species. This will be discussed in the results section of the report. Unlike the other methods used in this report, this is a non-hypothesis testing, exploratory data technique. However, it does help draw inferences.

3.0 Results

3.1 Differences among sites 2007 habitat

The sampling regime for all other variables was based on selecting samples from a uniform flow range between sites. This was a very important aspect of the sampling plan because it allowed us to stratify the sampling plan without bias (a statistical concern) and allowed us to control an unwanted source of variation on the invertebrate community. That is, invertebrate assemblages are known to respond to water velocity, but we do not anticipate development on or near the Pinedale Anticline to increase the velocity of water in the New Fork River. Thus, if we collected samples from the same approximate range of flows at all sites, we could control for this variation and account for it statistically.

For most of the samples, the field crew was able to sample from a uniform range of near stream-bottom flows (Fig. 3.1). However there were several deviations from this level of uniformity. First, at sites NF 30 and NF 40, samples were collected from lower flows on average from all other sites. Additionally, samples from NF60 were collected from a very wide range of flows. The confounding interactions in 2007 were limited, but the overall performance of the statistics would have been improved if consistent ranges were sampled at each site. Field crews should consider sampling a slightly wider range at all sites if it means they can find a uniform representation of habitats. Since natural particle assortments can be related to flow, and since the potential impacts of development within the PAPA can alter particle size distributions, it is imperative that field crews sample congruent flow regimes.

An Analysis of Variance (ANOVA) on the raw flow data indicated a significant difference among sites(P=0.033), but the follow up test (Tukey's HSD) to describe which sites were significant from each other, failed to describe any differences. This may be due to violation of assumption of homogeneity of variances among sites by site NF60 (Fig 3.1). When variances were homogenized by transforming data with natural logarithms, ANOVA did not indicate a significant difference among sites (P=0.148). This deviation in procedure could have caused some analytical problems in the terms of significant interactions. Fortunately our analyses were sufficiently robust this year to avoid these problems. Field crews need to be judicious in their collection of samples from congruent flow regimes in future assessments.

.

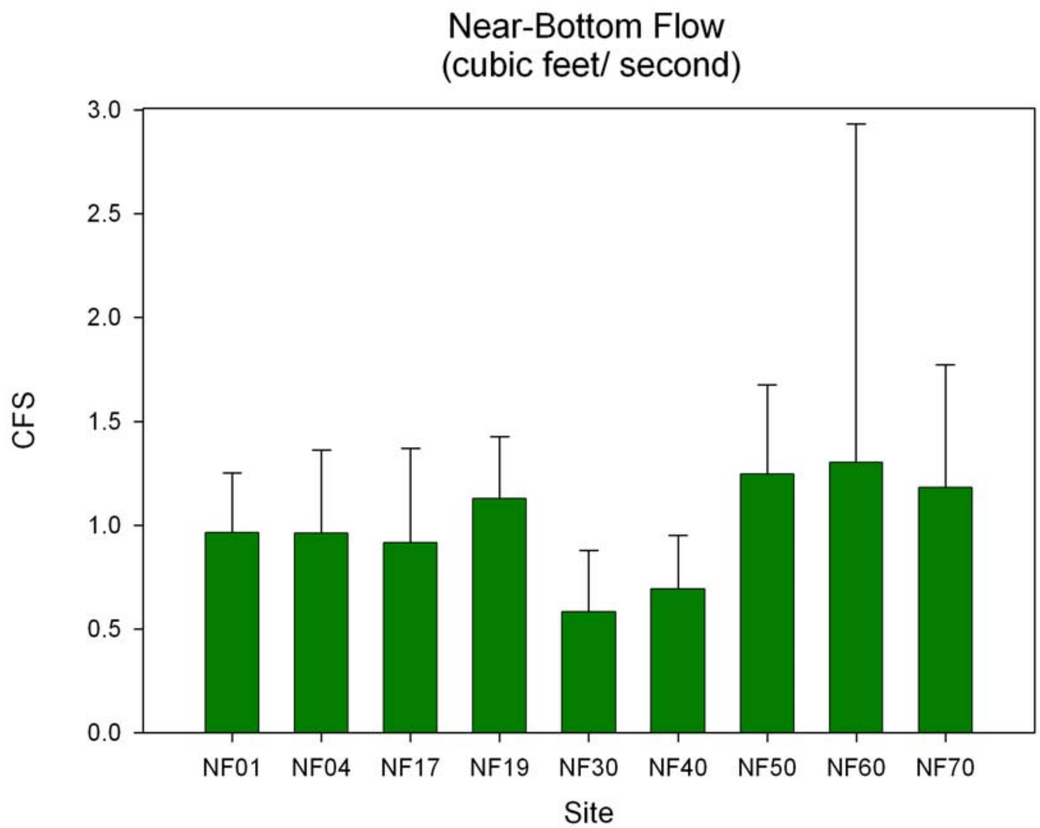


Figure 3.1. Near-Bottom Flow. The velocity of water was measured as close to the bottom of the river to attain a useful covariate for macroinvertebrate assemblages. For this study, it would be ideal for the bars to be the same height and the errors-bars to about the same width. (bars are means, with 95% CI). Through out this report, measurements made near the river-bottom are called "near substrata" measurements.

The size of particles comprising the stream bottom is important for the success of macroinvertebrates as well as for fish reproduction. The field crew quantified the size composition substrata (Fig 3.2) within each Surber sample. Thus, these data do not describe the totality of particle sizes found at each site, but rather where the benthic samples were collected. Since the effort was standardized by water velocity, the sites should be somewhat similar—unless something other than flow has influenced the distribution of particles in the river. For example, we know from field observations, that smaller particles should naturally dominate the East Fork River (NF17).

We found that there was more sand in the samples collected from NF30 and NF40. The levels at these sites exceeded the proportion of sand in samples from the East Fork River (NF17), which is sand-dominated. This may be a symptom of PAPA related sedimentation or it may be related to the fact that slightly lower flows were sampled at NF30 andNF40 than at NF17 (and the other sites).

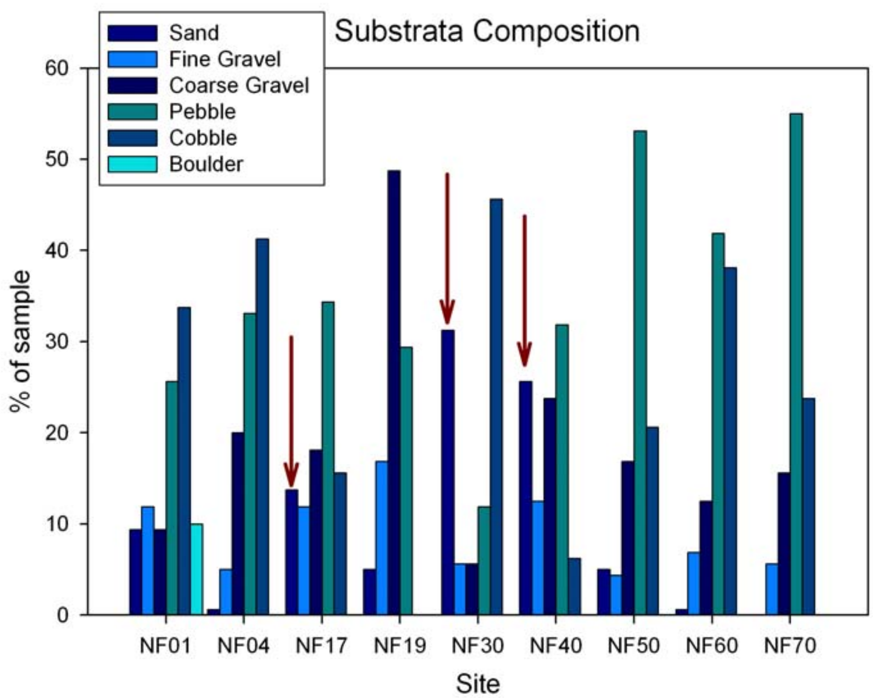


Figure 3.2. Substrata Size Composition. The size of stones in the individual Surber samplers was characterized using the Wentworth Size scale—which is standard for invertebrate studies (e.g., Minshall 1984). The bars are means, but error bars are not displayed so that disparities in the composition of different size classes can be observed.

The question of causation could not be addressed using composited Surber samples. But we were able to examine the relationship in greater detail using the single samples. We know that depositional areas[17] accumulate fine particles because of reduced water velocity. At higher flows, sand and silt are entrained into the water column and swept down stream until they reach a lower flow and are deposited on the river bottom—much like snow drifts and their relation to hillsides, or snow fences. Thus, a natural relationship would be one in which sand deposition would decrease as near-substrate water velocity increases. Deviations from this relationship indicate that something different is occurring.

In the case of the elevated sand levels at NF30 and NF40, we have observed more sand among the stone substrata, but we also sampled at a lower flow. The way to determine if there is a difference in the sand-flow relationship among the sites is to process the data by least squares regression; differences in the relationship will be represented as differences in the slope of the lines. When we performed this analysis, we found that there was no significant difference in the slopes of NF17 or NF40 but that NF30 actually showed an inverse relationship—where flow water velocity seemed to elevate the concentrations of sand in the samples. This is consistent with an actively eroding system, and could be caused by activity near the pipeline crossing between NF30 and NF40. The question remains, was this sufficient to influence the biota of the New Fork River?

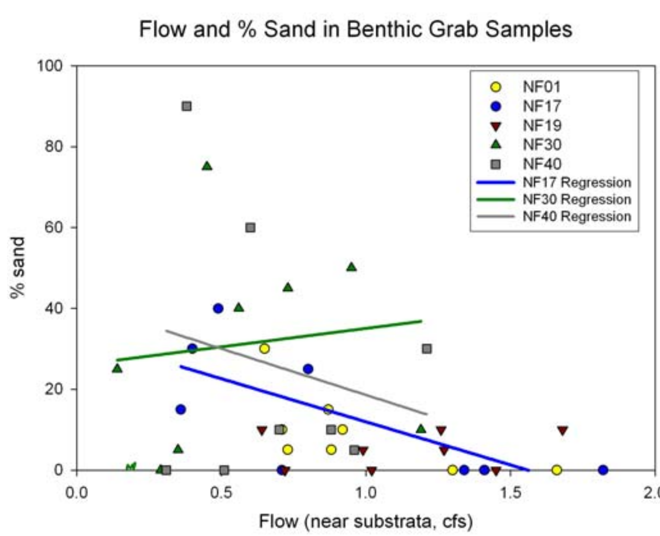


Figure 3.3. Regression of Velocity and % Sand. The relationship should be negative; elevated water velocity should result in decreased sand deposition among the substrata. This held true for NF17 and NF40, but not for NF30, which had a weak, positive relationship between flow and sand deposition in the samples. Note that even though the relationship was inverse, the relationship is sufficiently weak that it was not statistically significant when tested (ANCOVA: P=0.287).

---

[17] Aquatic biologists consider places where finer particles to fall (or deposit) from water column and accumulate on the stream bottom as depositional habitats. This is different from the definition used by geomorphologists.

3.2 Differences among sites 2007 biological metrics

Overall differences among sites

The biological metrics were first screened for significant differences among sites using ANOVA (Table 3.1). Metrics which displayed a significant difference among sites were examined in greater detail to determine if the effects could be due to impairment related to development in the PAPA. Many of the 15 metrics tested indicated that there were statistically significant differences among the sites (P<0.05, Table 3.1).

TABLE 3.1. ANOVA Results. The ANOVA resulted in statistically significant differences among sites for 10 of the 15 metrics tested directly. For metrics that exhibited significant differences, Tukey's HSD was used to identify which sites were significantly different from each other. Sites that were significantly different from each other are noted by different letters in their columns. Sites were not significantly different from each other share at least one letter with similar sites.

			MULTIPLE COMPARISONS GROUPING (TUKEYS HSD)								
Metric	P-val		NF01	NF04	NF70	NF60	NF17	NF50	NF40	NF30	NF19
Taxa Richness	0.115		--	--	--	--	--	--	--	--	--
E-Taxa	0.282		--	--	--	--	--	--	--	--	--
P-Taxa	<0.001		A	B	AC	AC	BC	BC	C	B	AC
T-Taxa	0.095	*	A	AB	AB	A	A	AB	A	AB	B
% E	0.109		--	--	--	--	--	--	--	--	--
% P	<0.001		A	B	B	B	B	B	B	B	B
% T	<0.001		A	A	C	AB	B	AB	C	BC	C
%Chironomid	<0.001		D	AC	B	C	A	AC	B	B	B
%Non-Insect	<0.001		A	A	C	C	AC	A	AC	B	A
HBI	<0.001		A	B	A	B	C	C	C	C	B
Gatherers	<0.001		AC	C	B	B	B	B	B	A	B
Filterers	0.001		BC	A	B	B	C	B	BC	C	BC
Scrapers	<0.001		A	BC	B	B	BC	BC	AC	AC	B
Semivoltine	0.002		A	B	B	B	B	B	B	B	B
Dominance(5)	0.240		--	--	--	--	--	--	--	--	--

Taxa Richness

The initial GLM indicated that there were no significant differences among the sites based on the taxonomic richness of SS samples (Table 3.1). The metric increased with particle size index (P=0.002) and had marginal correlation to the interaction of particle size and water velocity (P=0.081). Overall the differences in the number of species sampled differed very little among sites.

Table 3.2. Taxa Richness ANOVA. There was no significant difference among the taxa richness exhibited by sites sampled using SS methods.

Dep Var: TAXA N: 72 Multiple R: 0.420 Squared multiple R: 0.177

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	382.000	8	47.750	1.690	0.118
Error	780.000	63	28.254		

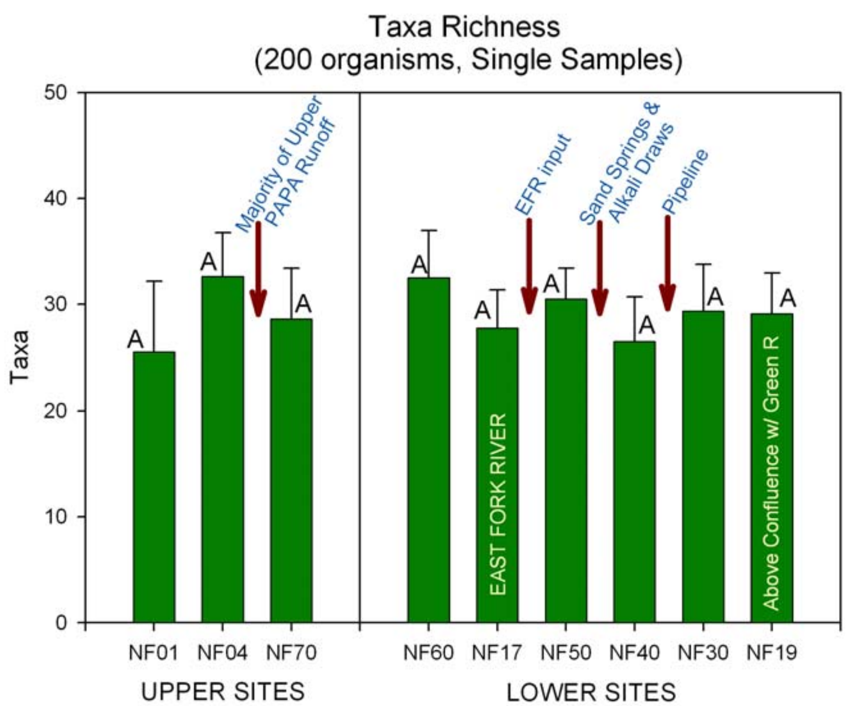


Figure 3.4. Taxa Richness. There was no significant difference among the taxa richness exhibited by sites sampled using SS methods

Plecoptera Richness

Statistically, there was much variation in the number of stonefly species occurring at the sites, when sampled by the SS methods(Table 3.1; P<0.001). It is important to keep the scale of these differences in mind. The site with the greatest richness of Plecoptera taxa was NF01, which had an average of about 2.7 species, and included some smaller stream nemourid species. Most of the other sites averaged about 0.8- 1.6 species (Fig 3.5). Thus the actual scale of these statistically significant results is quite small. The metric was significantly influenced by particle size (P=0.039) and marginally significantly influenced by flow (P=0.051). When the flow-corrected means were tested through ANCOVA, the differences altered only slightly (P=0.001; Fig 3.6) .

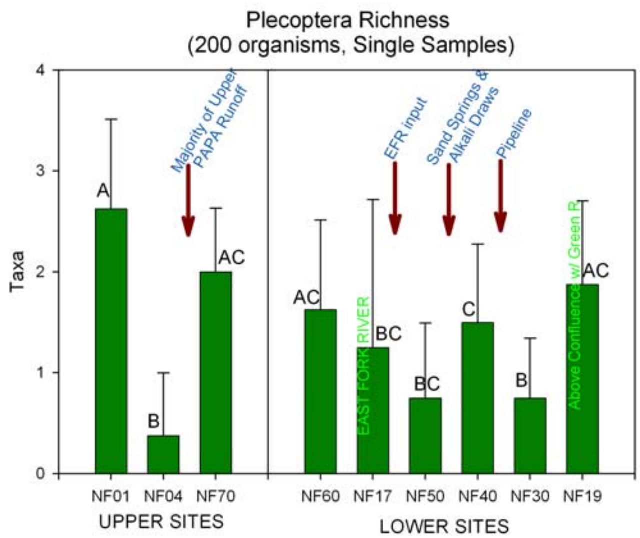


Figure 3.5. Plecoptera Richness. The raw richness of Plecoptera is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked "BC" is not significantly different other bars marked with either a "B" or a "C," but is significantly different from sites marked "A."

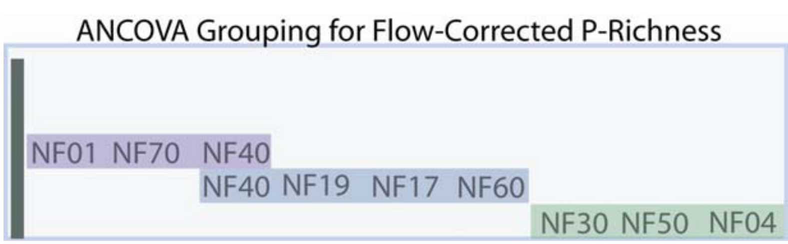


Figure 3.6. Flow-corrected Plecoptera Richness. The statistical grouping for Plecoptera Richness after adjusting for the influence of near-substrate water flow. Sites on the same row are not significantly different from each other, sites listed on higher rows had significantly higher average values for the metric than those on lower rows.

Trichoptera Richness

Only marginally statistically significant differences were observed among the sites' Trichoptera richness (Table 3.1; P<0.095). These differences were caused primarily by the difference between NF01 and NF19—no sites between them were significantly different from each other (Fig. 3.7). This reflects the change in richness between the farthest upstream site and the farthest down stream sites. Moreover, the downstream site supported more species than the farthest upstream site—which is not consistent with ecological perturbation. The metric was influenced mostly by particle size and when the variation associated with particle size was accounted for, differences among sites no longer explained significant amounts of variation (P=0.121). Marginally significant differences among sites persisted after flow adjustment (P=0.062) when particle size was not included in the model. Since this metric did respond to particle size index, it is likely to respond to sedimentation effects over the long-term.

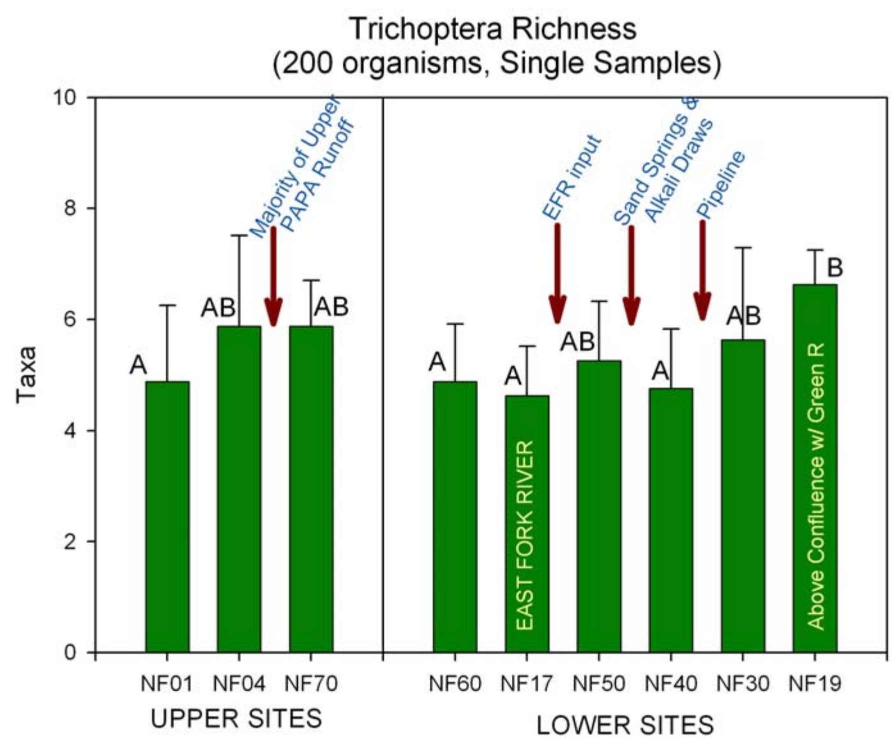


Figure 3.7. Trichoptera Richness. The richness of Trichoptera is presented (± 95%CI) with multiple comparison[18] results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked "BC" is not significantly different other bars marked with either a "B" or a "C," but is significantly different from sites marked "A."

[18] Multiple or pairwise comparisons are statistical tests that follow ANOVA to describe which specific sites are different from each other.

Plecoptera Relative Abundance

Statistically, there was much variation in the relative abundance of stoneflies occurring at the sites, when sampled by the SS methods(Table 3.1; P<0.001). However this was caused by the large population of nemourid stoneflies (such as *Zapada cinctipes)* which usually occur in smaller shaded cool streams. None of the other sites exhibited statistically significant differences in the abundance of Plecoptera which were uniformly uncommon—generally averaging less than 3% of the population (Fig. 3.8). The difference observed is believed to be related to natural differences between the farthest upstream site and the others. If the metric were to respond to development on the PAPA, we would have expected to observe significant differences among more sites.

This metric was responsive to both % sand (P=0.005) and the particle size index (P=0.008), but since relatively large differences existed between NF01 and the other sites, differences among sites remained significant (P<0.001) even after correction for the influence of these variables by ANCOVA.

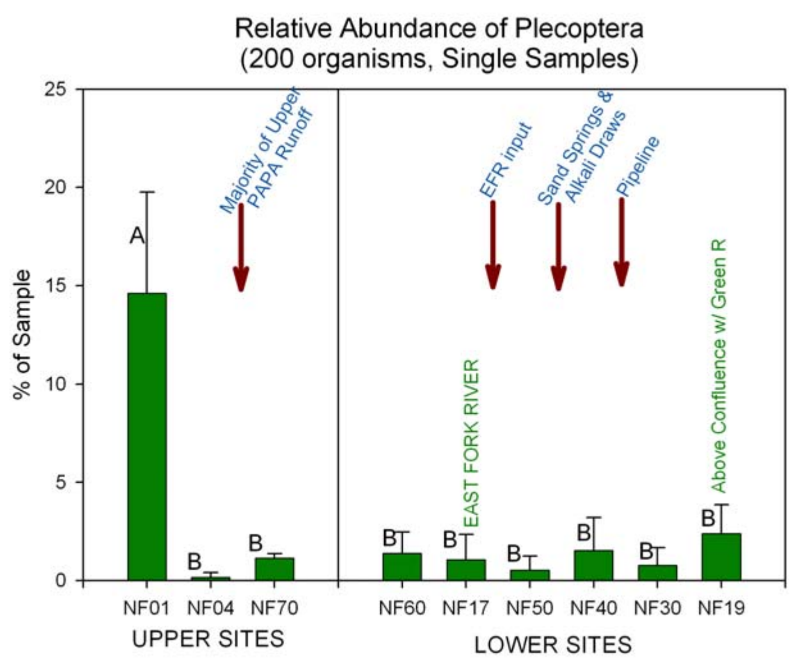


Figure 3.8. Plecoptera Abundance. The relative abundance of Plecoptera is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked "BC" is not significantly different from other bars marked with either a "B" or a "C," but is significantly different from sites marked "A."

Trichoptera Relative Abundance

Statistically, there was much variation in the relative abundance of caddisflies occurring at the sites, when sampled by the SS methods(Table 3.1; $P<0.001$). The upstream sites (NF01 and NF04) produced samples containing fewer Trichoptera than other sites (Fig. 3.9). The differences were largely due to the addition of the taxon *Hydroptila* sp. in samples from NF17 and below (Fig. 3.10)—except for NF50. Hydroptila may have responded to sand downstream, or the filamentous algae upon which it usually feeds. There was no algae data to correlate with *Hydroptila* abundance data, but since the taxon was abundant in the East Fork River (NF17), which we assume is largely unaffected by development in the PAPA, it is unlikely that the increase in this taxon was related to activities in the PAPA.

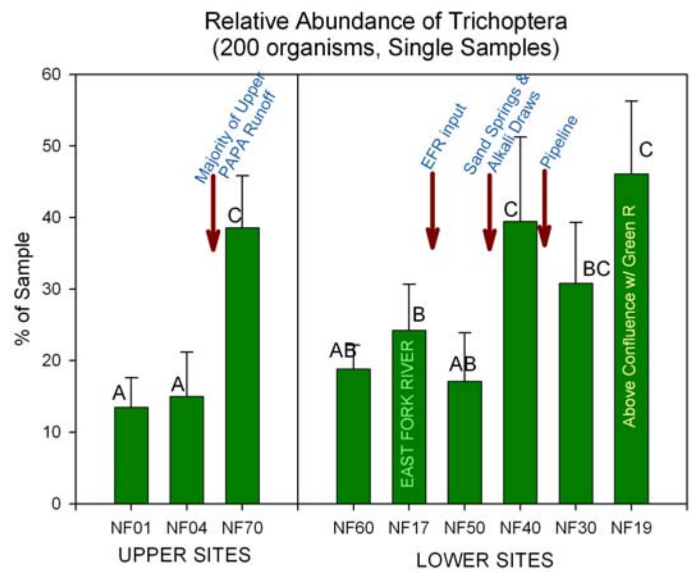
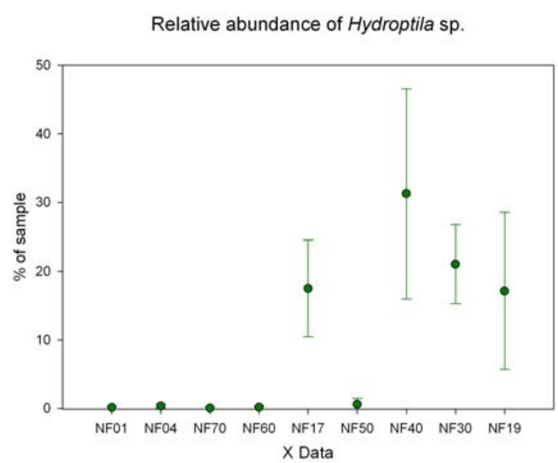


Figure 3.9. Trichoptera Abundance. The relative abundance of Trichoptera is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked "BC" is not significantly different from other bars marked with either a "B" or a "C," but is significantly different from sites marked "A."

Figure 3.10. *Hydroptila* abundance. The relative abundance of "micro-caddisflies," *Hydroptila* sp. increased dramatically downstream from the confluence of the East Fork River. (means ± 95%CI)



Chironomidae Relative Abundance

The lower study area showed an increase in the abundance of midges downstream from the East Fork River—which had the greatest proportion of midges of all locations (Fig. 3.11). Samples from NF01 had the least contribution of chironomid midges of all sites, and most of the sites between NF01 and NF19 were not statistically different from each other. The exception was from NF17, which had the greatest midge dominance and was not significantly different from NF50 or NF40. It is important to realize that there are about 4000 species of midges in the northern hemisphere and that some are very sensitive to disturbance. However, when midges dominate a river system, especially a smaller river like the New Fork, it is usually considered a sign of stress or disturbance. Most of the sites below the PAPA development influences showed a reduction in the dominance of midges, indicating that this metric was not responding to PAPA related development in 2007.

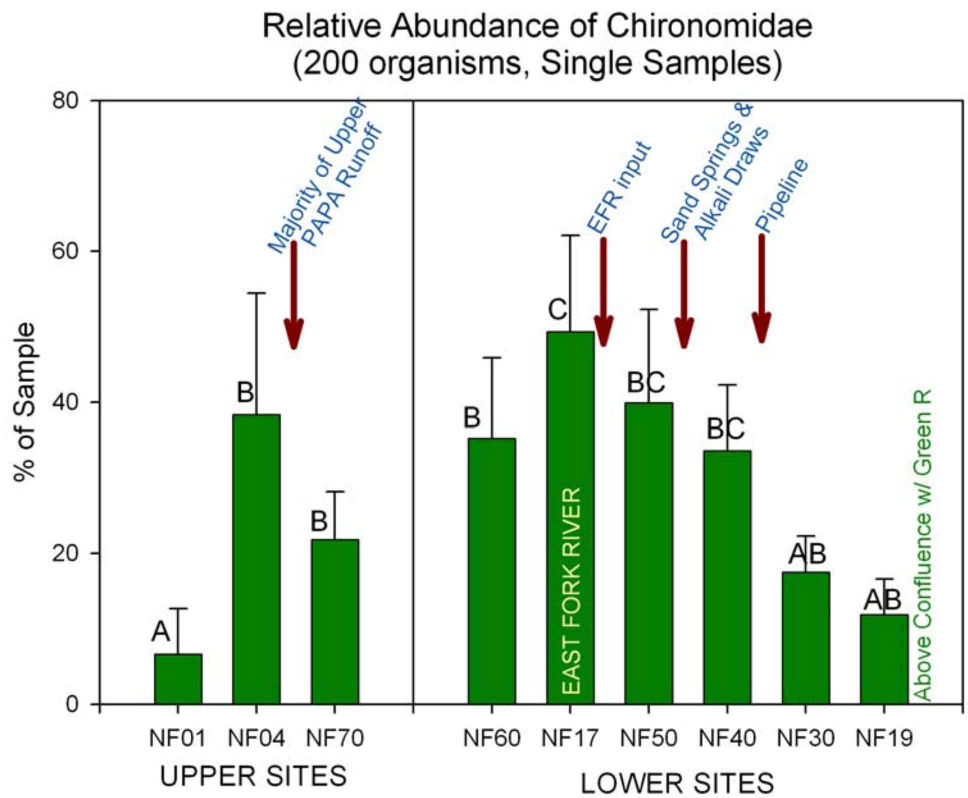


Figure 3.11. Chironomidae Abundance. The relative abundance of Chironomidae is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked "BC"is not significantly different from other bars marked with either a "B" or a "C," but is significantly different from sites marked "A."

Non-Insect Relative Abundance

Like the Chironomidae, there are both sensitive and tolerant species of non-insects. But these usually only comprise a small portion of the sample (>10%) in “normal” North American Rivers. The lowest dominance on non-insects occurred at NF70 and the next site downriver, NF60 (Fig. 3.12). Most of the sites were not significantly different from each other. The samples collected from NF30 contained significantly more non-insects than all other sites. This has been part of a larger trend at the site since 2004 (Fig. 3.13). As with previous years, when this site had elevated non-insect abundance, it was largely due to increased abundance of aquatic oligochaete worms.

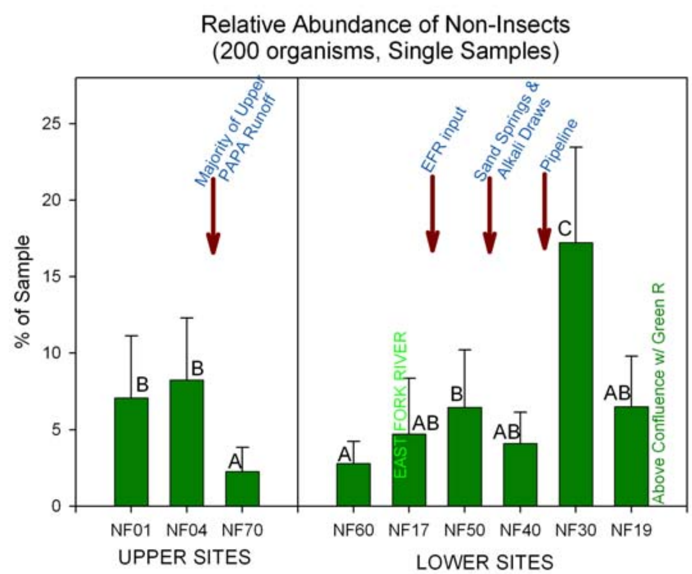


Figure 3.11.Non-insect Abundance. The relative abundance of non-insects is presented (± 95%CI) with multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. For example, a site marked “BC“ is not significantly different other bars marked with either a “B” or a “C,” but is significantly different from sites marked “A.”

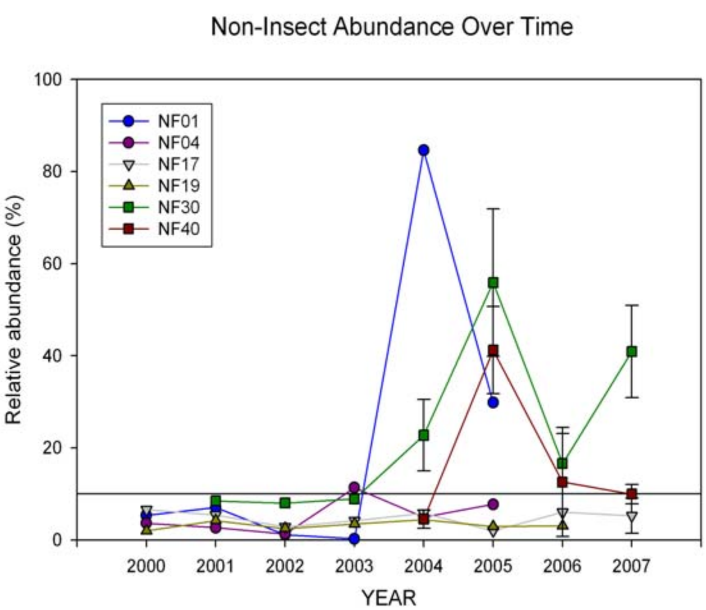


Figure 3.12. Non-insect Abundance over Time. The relative abundance of non-insects in SS samples is presented (± 95%CI) since the beginning of baseline biological monitoring on the New Fork River. Elevated non-insect abundance began in 2004.

River Continuum Concepts

Hilsenhoff Biotic Index

The Hilsenhoff biotic Index weights samples by the abundance of organisms and their tolerance to organic pollution. The HBI varies from 0 to 10, with zeros indicating a community that is very sensitive to pollution and 10 is indicative of communities dominated by sludge-dwelling organisms.

The metric seemed to respond to the confluence of the East Fork River and remained elevated until the farthest downstream site, NF19, before the confluence with the Green River. The metric was not responsive to %sand (P=0.972), embeddedness (P=0.182) or particle size (P=0.436). Flow did explain a marginal amount of variation in the metric (P=0.097), but this did not change the significance of the differences among the sites in 2007. This metric did not appear to respond to PAPA development.

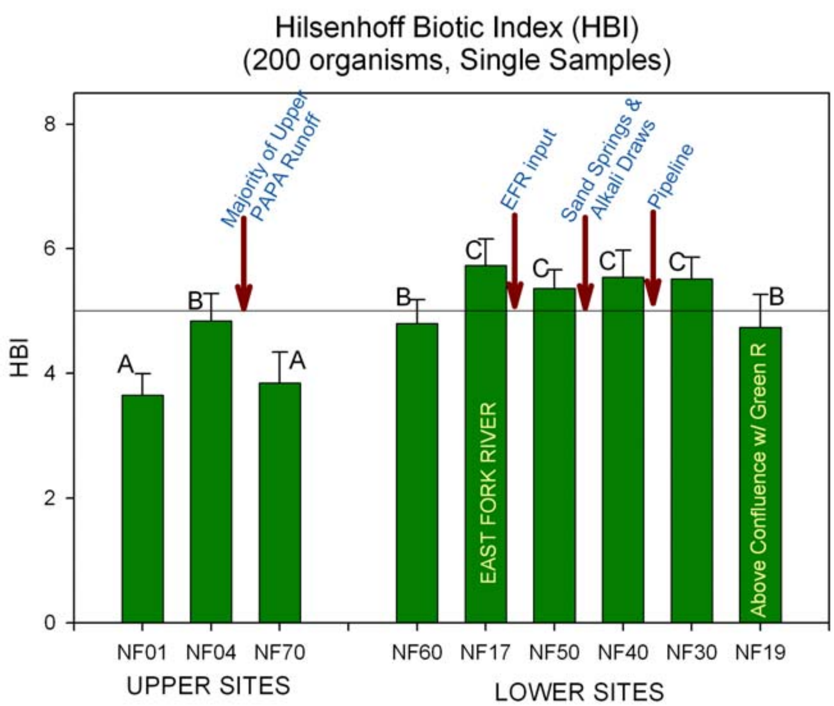


Figure 3.13. Hilsenhoff Biotic Index (HBI). The HBI is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. The line marks HBI = 5.0; above this value most of the organisms are moderately (or more) tolerant to organic pollution. Values below 5.0 occur when most the organisms are less than moderately tolerant to organic pollution.

Abundance of Collector-Gatherers

Collector-gatherers are generalists that feed on fine pieces of organic material. Their abundance usually increases after disturbances because specialists are often displaced by disturbance. Site NF30 had significantly more Collector-Gatherers than all sites except NF01 (Fig 3.14). The abundance of Collector-gatherers at NF30 was a direct result of elevated non-insects because most of them were collector-gathering worms (Figs. 11, 12). This measure may be evidence of some moderate, localized PAPA-related disturbance. This metric did not correlate significantly with any environmental covariates.

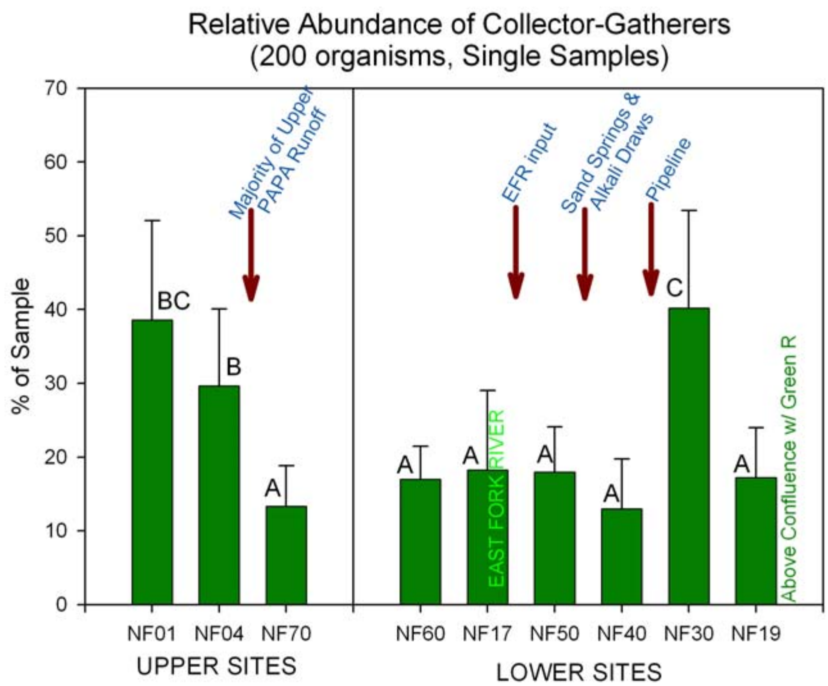


Figure 3.14. Relative Abundance of Collector-Gatherers. Collector-gatherer abundance is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other.

Abundance of Collector-Filterers

Collector-filterers are similar to collector gatherers, except that they consume organic particles that are suspended in the water. Thus, they usually respond to changes in the amount of suspended organic material (e.g., slight sewage influences) by increasing their abundance.

Site NF04 had significantly greater abundance of collector gatherers than all other sites. Two other sites, NF17 and NF30, had significantly lower abundance of collector-filterers than the other sites (Fig. 3.15). For NF30, this was because of the elevated abundance of collector-gatherers (aquatic worms), and macrophyte piercers (i.e., *Hydroptila* sp.). NF17 offset lower collector-filterers with macrophyte piercers and Shredders[19]. The metric was correlated with flow (P=0.016) and after correction for the influence of flow, the only difference that remained statistically significant was NF04 from all other sites (Fig. 3.16).

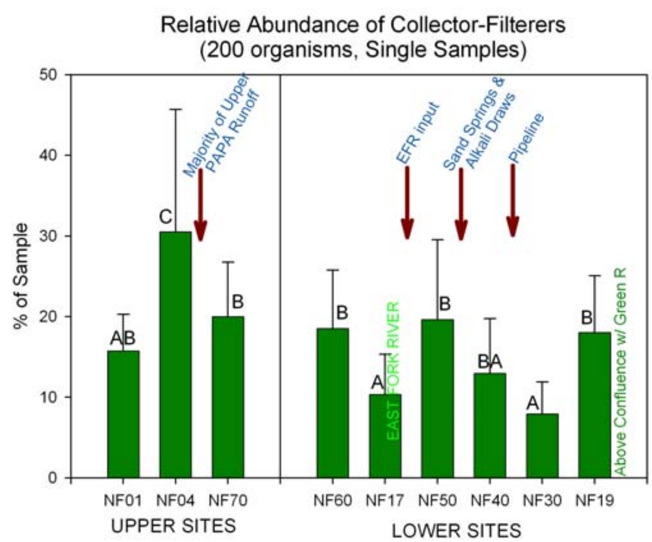


Figure 3.15. Relative Abundance of Collector-Filterers. Collector-filterer abundance is presented (± 95%CI) with multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other.

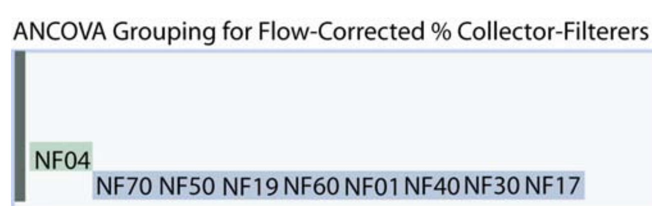


Figure 3.16. Flow-Corrected Abundance of Collector-Filterers. The statistical grouping for collector-filterer abundance after adjusting for the influence of near-substrate water flow. Sites on the same row are not significantly different from each other, sites listed on higher rows had significantly higher average values for the metric than those on lower rows.

[19] "Shredders" were not typical shredder-detritivore functional feeding group. The taxonomy lab assigned shredders to *Cricotopus nostococladius* which is an obligate symbiant with the blue-green algae *Nostoc* sp. Thus it is related to shredder-herbivores.

Semivoltine Richness

Semivoltine animals are organisms which take more than one-year to reach sexual maturity. The rationale for this as a metric is that semivoltine animals are usually poor colonizers. Most of the semivoltine taxa collected in this study were stoneflies. Thus it makes sense that the richness of Semivoltine taxa very much resembles the richness of Plecoptera (above). The primary difference observed was a difference in the richness of semivoltine taxa at the far-upstream site (NF01) from many of the downstream sites. Site NF19, the downstream site was marginally different from NF30 ($0.05 < P < 0.10$). As with Plecoptera richness, it is important to note the scale of the differences. The lowest sites had an average of one semivoltine taxon, and the highest had an average of 2.4 taxa. These differences are quite small.

This metric has been unresponsive in many assessments I have run because the true semivoltine taxa are usually low in abundance. We ran the analysis on this metric because it is one of the metrics used to construct the WSII.

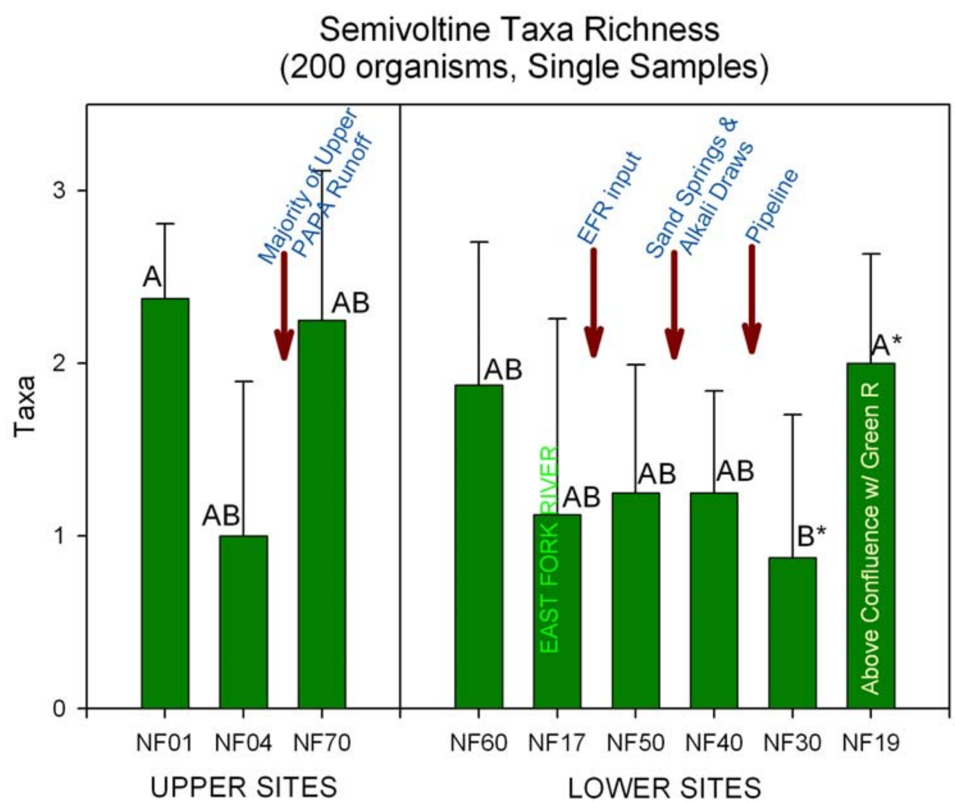


Figure 3.17. Semivoltine Taxa. Semivoltine Richness is presented (± 95%CI) with Multiple comparison results noting which sites were significantly different from each other. Bars with the same letter are not significantly different from each other. Semivoltine taxa are separated from long-lived taxa, like beetles (Hargett and Zumburge 2006). Most of the semivoltine taxa observed were stoneflies. The sites marked with an * were marginally significantly different from each other ($0.05 < P < 0.10$).

3.3 Differences among sites: Ordination of Species

In aquatic ecology, it is impractical to compare each species to each other for each and every site because not all species occur at all sites naturally. Also, the number of zero-values results in violations of the assumptions of many statistical techniques. Finally, the number of comparisons (like those in Table 3.1) would be very confusing. Imagine interpreting this table for ~200 species.

In 2007 we have individual samples with habitat variables associated with each one. This allows us to use a constrained ordination technique to examine how the samples differ in overall species composition, in response to environmental gradients. Ordinations are methods of multivariate statistics that project axes though an N-dimensional hyperspace defined by the number of species (~200 dimensions) and project these multi-dimensional axes as simple linear axes—allowing us to visualize trends that we could not visualize or test other ways. Ordination techniques are descriptive—there is no hypothesis test involved.

We used Detrended Correspondence Analysis (DCA) to prepare the data for Canonical Correspondence Analysis (CCA) to project these relationships, because the ordination is constrained by the habitat variables used, reducing the chance of spurious conclusions. However, all ordinations were so strongly influenced by the difference in NF01's community structure that it obscured trends among the other sites (Fig. 3.XX). This was partially due to the unique species that live at NF01, and partially due to the low abundance of invertebrates from that site. The most influential taxon in the analysis was the small stonefly *Zapada cinctipes* which was not collected anywhere else. This species usually lives in smaller, cooler, forested streams and eats decaying leaves that accumulate among the substrata. The New Fork River, over most of the study area is larger and less shaded than where this stonefly family typically occurs.

Because samples from NF01 represented an unattainable reference (that is the other sites are so fundamentally different that it is unlikely that the other sites would ever resemble NF01 even in the absence of PAPA development), we chose to re-run the analysis while excluding NF01.

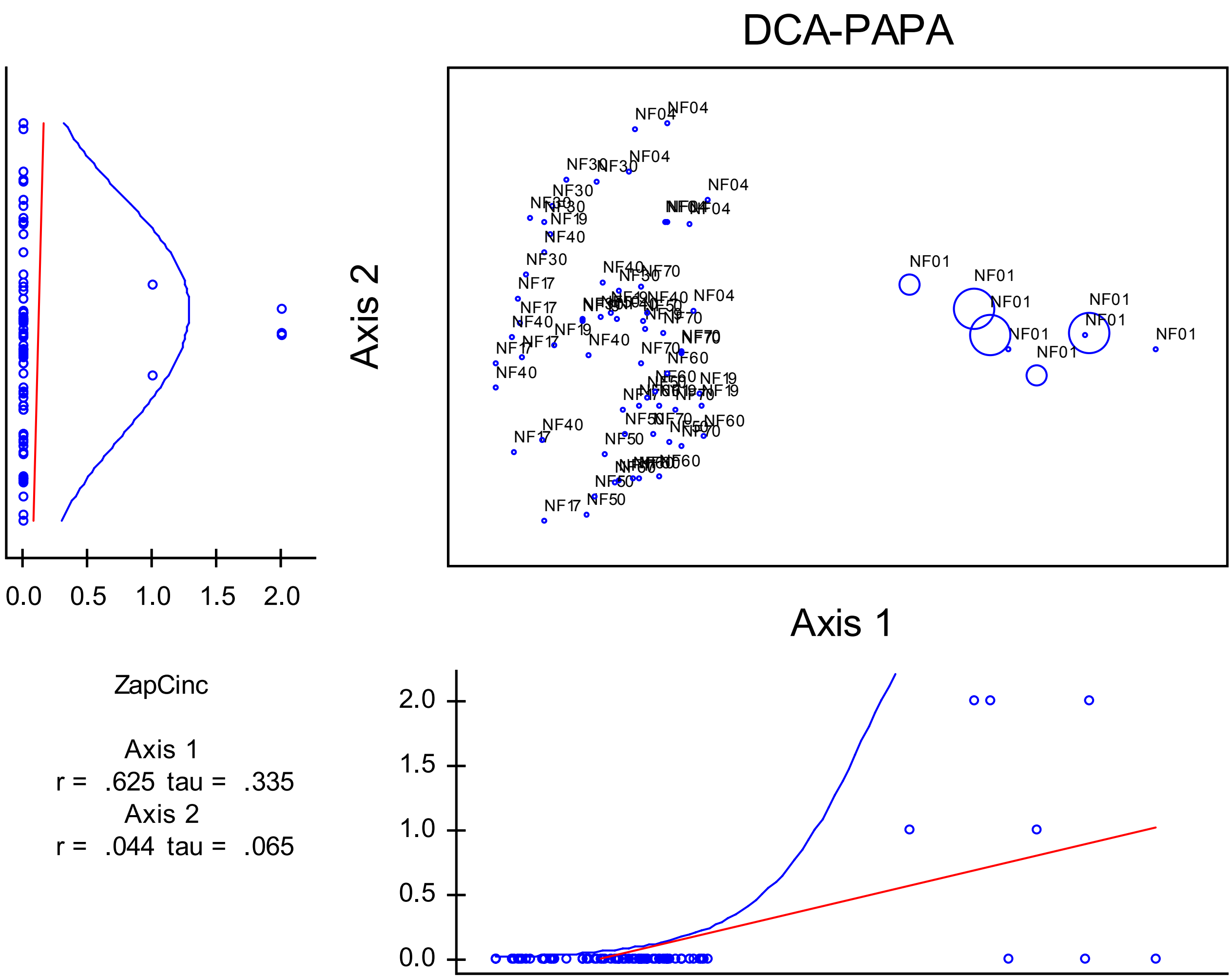


Fig. 3.18. Detrended Correspondence Analysis. The Detrended Correspondence Analysis for all sites was overwhelmingly driven by the unique taxonomic composition of samples collected from NF01. This was largely due to the occurrence of the tiny stonefly *Zapada cinctipes*, and the low abundance of organisms in the samples (individual Surber sample abundance ranged from 34-915, with three samples containing less than the 200 organism sorting target). The unique taxonomic structure of this site combined with low abundance of some samples made the site stand out from all others. In the main figure, the groupings of the sites are determined by the relative abundance of all species—and the size of the circles is proportional to the influence *Z. cinctipes*, which only occurred at NF01. The figure on the bottom shows how much influence these sites had on the left-to-right grouping of all samples (Axis-1 is multivariate combination of influential taxa). Thus the primary axis was driven almost entirely by the samples of NF01. We believe that these differences are due solely to the uniqueness of NF01 and not related to PAPA development. Therefore we excluded these sites from the multivariate analysis and repeated the procedure so that differences among the sites of interest could be discerned.

When the DCA was rerun without NF01, the primary and secondary DCA Axes were more appropriate because they were not influenced by taxa unique to NF01. Thus they reflect the actual differences occurring among the sites of interest rather than an unattainable reference (NF01). Here, the influence of a common species (*Baetis tricaudatus*) is shown (Fig. 3.19). This unconstrained DCA shows the relationship of the samples based on the inter-relationship of the species that comprised the samples. For example NF04 had a unique species composition, which was influenced by *Rheotanytarsus* (a filter-feeder) and several other midges (these taxa had a heavy positive loading on Axis-2) NF30, NF40, and NF17 had fewer *B. tricaudatus* than normal (it should be ubiquitous). Since the abundance of *B. tricaudatus* has a negative loading on Axis-1 (high abundance of this taxon pulls samples to the left), these sites came out largely on the right site side of the biplot (Fig. 3.19) This data set was used as the basis for the constrained CCA (presented next).

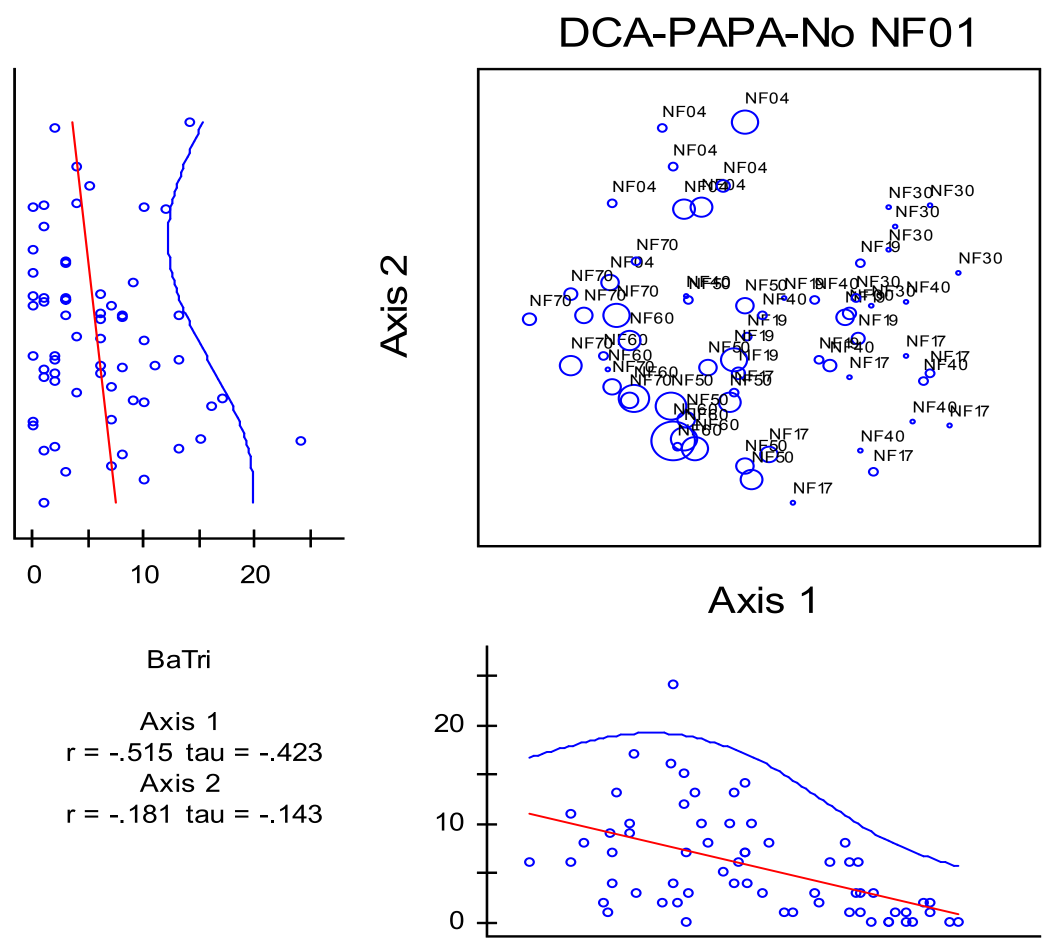


Figure 3.19. Detrended Correspondence Analysis. The Detrended Correspondence Analysis excluding NF01. Trends are different and clearer than they were with the inclusion of NF01. Note that sites NF40 and NF17 had samples in the lower right region. Sites NF50 and NF60 were near the center or lower left. NF30 was to the upper right and NF04 was to the upper left. The side graphs show the factor loading for *Baetis tricaudatus*, a mayfly that should occur at all sites. samples on the right side of the graph usually had very low abundances of this common species—which is noted by the size of the circle in the main DCA Plot.

The CCA indicated that the overall taxonomic structure of the samples moves the plot of samples to the right when they are influenced by high levels of sand. Samples are pulled to left if they are dominated by larger particles, and faster water. Vertical distribution of samples is more complex. Samples are pulled down and to the right when embeddedness is high, and pushed up and to the left when it is low. Gravel pushes samples up and to the right. The size of the circles is relative to the influence of Oligochaete on the ordination (Fig. 3.20).

The results indicate that NF 30 and NF40 were both strongly influenced by the occurrence of sand and embeddedness which correlated with an abundance of oligochaete worms in most cases. NF30 samples tended to be along the periphery of the plot. These results corroborate the other finding earlier in the report—suggesting that NF30 (and perhaps NF40) is influenced by fine sediments—more so than NF17.

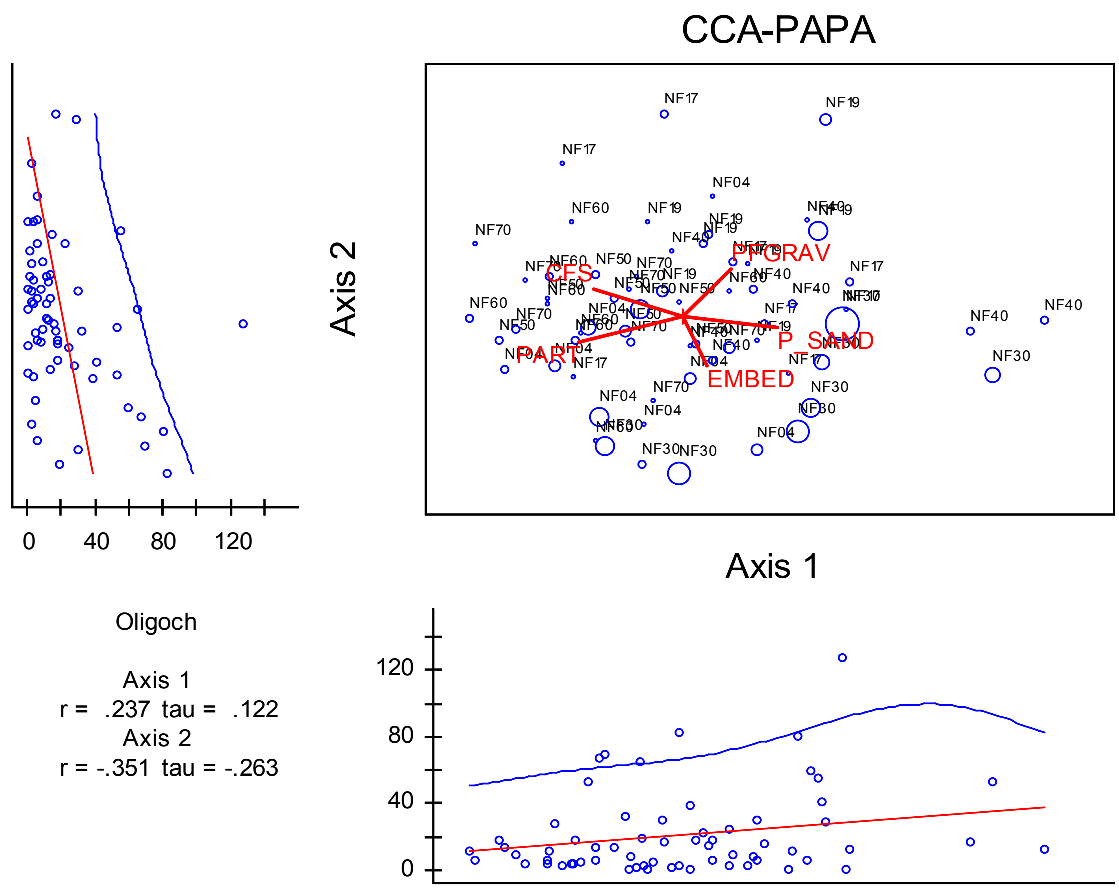


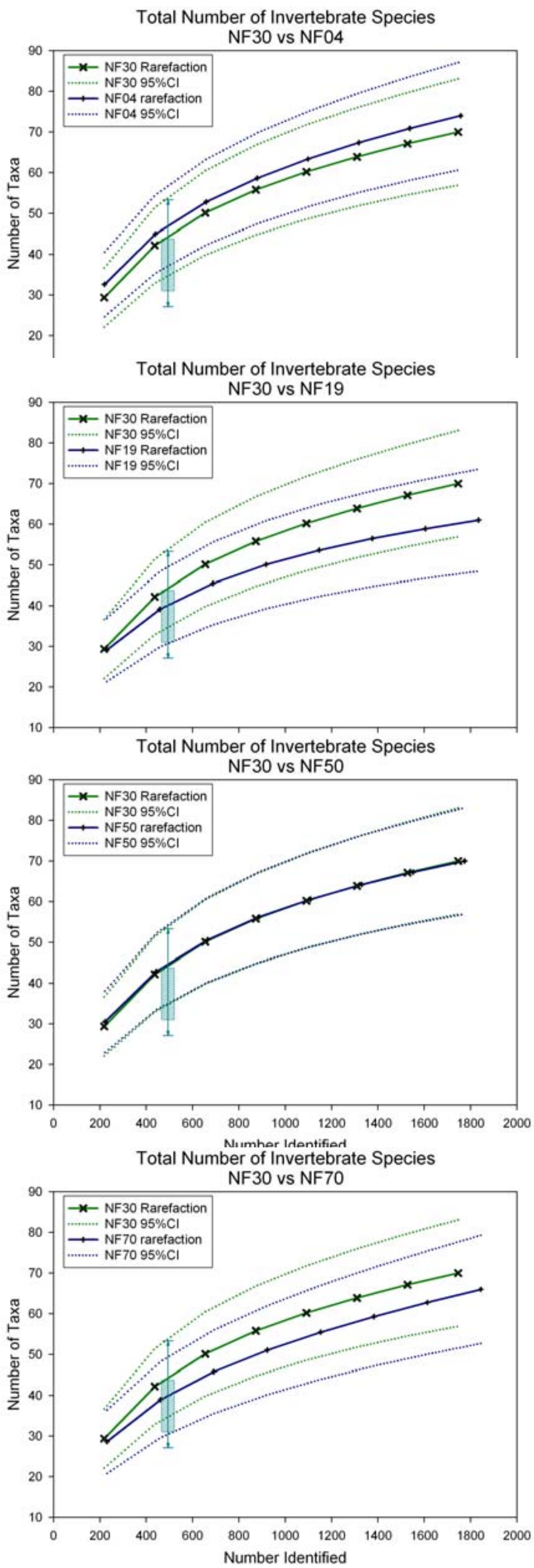
Figure 3.20. Canonical Correspondence Analysis. The Canonical Correspondence Analysis constrained the DCA-data for correlations with the environmental matrix. The side graphs show the factor loading for oligochaete worms which should be a very small constituent of samples (larger circles have large oligochaete populations). Also, the red vectors indicate the influence of environmental variables. Sand and embeddedness pulled samples to the right and down. This separated NF30 and some samples from NF40 from the other sites.

3.4 Differences in SS vs. CS WSII Results

We intend to use 2007 and 2008 data to evaluate the effects of different sampling techniques on the results generated by the Wyoming Stream Invertebrate Index (WSII) because the findings will be more useful if two or more years are tested. However, to complete this report we needed to know if the collection method had any significant difference on the WSII so that we could use the method that provided the greatest amount of information when compared to our historic data set. That is, if there is no significant difference in the methods, we may as well use the Single Sample (SS) data set so the results can be related to the rest of the 2007 assessment. Conversely, if there is a difference, we need to use the Composite Sample (CS) dataset to ensure the methods were compatible over time.

There is a relationship between the number of specimens identified and the number of species identified. For example, if you identify only one specimen, you can only identify one species. If you identify two specimens, you can only identify one or two species. If you identify three specimens, you could identify three, two or one species. This relationship defines a "species" accumulation curve. These are specific to the ecosystem studied and are the reason for standardized laboratory effort; if someone identifies 500 organisms, they may find a different number of species than someone who identifies only 100 or someone who examines 1000 specimens from the same habitat. There is a procedure called rarefaction analysis that allows you to back calculate the number of species you should have identified if you reduced the number of organisms you identified. Rarefaction analysis uses the number of organisms found from each species and simulates an electronic re-sampling of the population many times to arrive at an estimated species accumulation curve.

For all the richness measures used in the WSII (WSII-1 and WSII-2) we ran a rarefaction analysis to predict how many species we would have identified if we only subsampled 500 organisms—when in fact we identified about 1,600 organisms per site when the data were combined. Our combined eight individual Surber samples contained nearly 70 taxa (the WSII results found that most reference streams in the Wyoming Basin Ecoregion had fewer than 40 taxa when subsampled to 500 organisms). Thus, we used rarefaction analysis to project our approximate 70 taxa (depending on site) down to the actual number we would have found if we had only used 500 organisms. This was also done with other richness measures in the WSII (e.g., Ephemeroptera richness etc.) these numbers were used to calculate the WSII for the individual samples, which themselves only used ~200 identified organisms.



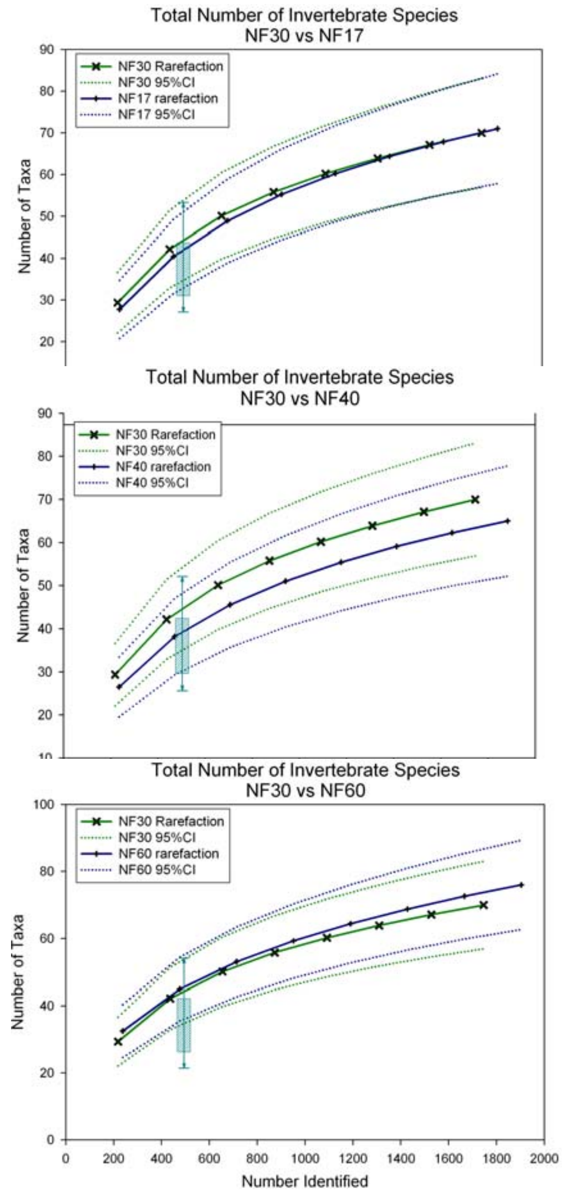


Figure 3.21. Rarefaction Analysis. When the eight single samples were electronically composited, Rarefaction Analysis was used to artificially sample the assemblage down to the 500-organism standard used by DEQ to calibrate the WSII. The large box plots displayed at 500-organisms denote the range of values observed by Stribling et al. (2000) for reference streams in the Wyoming Basin while calibrating the original WSII.

The Two-way ANOVA produced different findings for the old and new versions of the WSII (WSII-1, WSII-2, respectively). The WSII-1 resulted in a significant difference between NF30 and NF40, but no significant difference between methods (single sample vs. composite sample) and no significant interaction between the two treatments (Table 3.3).

The WSII-2 indicated that there were no differences between NF30 and NF40, but that there was a significant difference in the condition rating derived by the two methods (single sample vs. composite sample) even after the richness measures were adjusted by rarefaction (Table 3.4).

These findings were sufficient to warrant using the replicate composite samples for the comparisons of the sites over time. This has no effect on the findings reported earlier—which used raw metrics since the metrics were all calculated using uniform field and laboratory efforts. Next year's report will refine the methods for comparison so that we can eventually eliminate the large composite samples from the monitoring program without reducing the value of the baseline. However, we have insufficient data this year, and there were some confounded samples from the inconsistent flow sampling in 2007: to base our decisions on this year's data alone could produce spurious results.

Table 3.3. ANOVA WSII-1.

Dep Var: WSII-1 N: 26 Multiple R: 0.542 Squared multiple R: 0.294

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	326.502	1	326.502	6.123	0.022
DEVICE$	156.541	1	156.541	2.936	0.101
SITE$*DEVICE$	40.318	1	40.318	0.756	0.394
Error	1173.114	22	53.323		

Table 3.4. ANOVA WSII-2.

Dep Var: WSII-2 N: 26 Multiple R: 0.533 Squared multiple R: 0.284

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	6.321	1	6.321	0.224	0.640
DEVICE$	218.557	1	218.557	7.756	0.011
SITE$*DEVICE$	25.430	1	25.430	0.902	0.352
Error	619.975	22	28.181		

WSII Trends: The Composite Samples (CS)

Until we calibrate the single samples for comparability with the long-term dataset, we need to use the replicated composite samples to evaluate long term trends. Since we focused on differences among sites using the replicated single samples (SS), we limited the trend analysis to the two versions of the WSII (WSII-1, Stribling et al. 2000; and WSII-2, Harget and Zumberge 2006). Recall that both of these indices provided conflicting information about the condition of the sites in 2007 (above). WSII-1 indicated that there was a significant difference between NF30 and NF40, but no difference in the composite-sample or single-sample methods. However, the CS and SS methods produced significantly different WSII-2 scores even though there was no significant difference among the sites (above).

The ANOVA of these two methods over the entire monitoring program (using CS, 5 replicates) indicated that there were no significant differences over time (Tables 3.5, 3.6; Figs. 3.22, 3.23).

Table 3.5. ANOVA WSII-1 (CS).

Dep Var: WSII1 N: 40 Multiple R: 0.137 Squared multiple R: 0.019

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	193.148	1	193.148	0.052	0.821
YEAR	1603.845	3	534.615	0.143	0.933
SITE$*YEAR	570.145	3	190.048	0.051	0.985
Error	123232.727	33	3734.325		

Table 3.6. ANOVA WSII-2 (CS).

Dep Var: WSII2 N: 40 Multiple R: 0.115 Squared multiple R: 0.013

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	236.761	1	236.761	0.075	0.786
YEAR	805.194	3	268.398	0.085	0.968
SITE$*YEAR	361.984	3	120.661	0.038	0.990
Error	104688.692	33	3172.385		

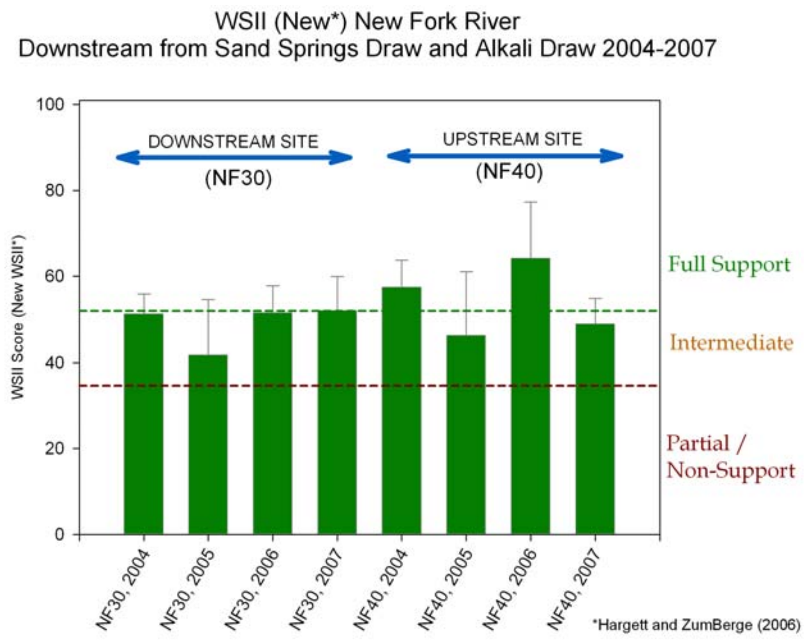


Figure 3.22. WSII-2. The composite samples indicated that there was no significant difference in the WSII-2 among years or between the sites NF30 and NF40. The sites changed very little and remained in the "intermediate" condition category.

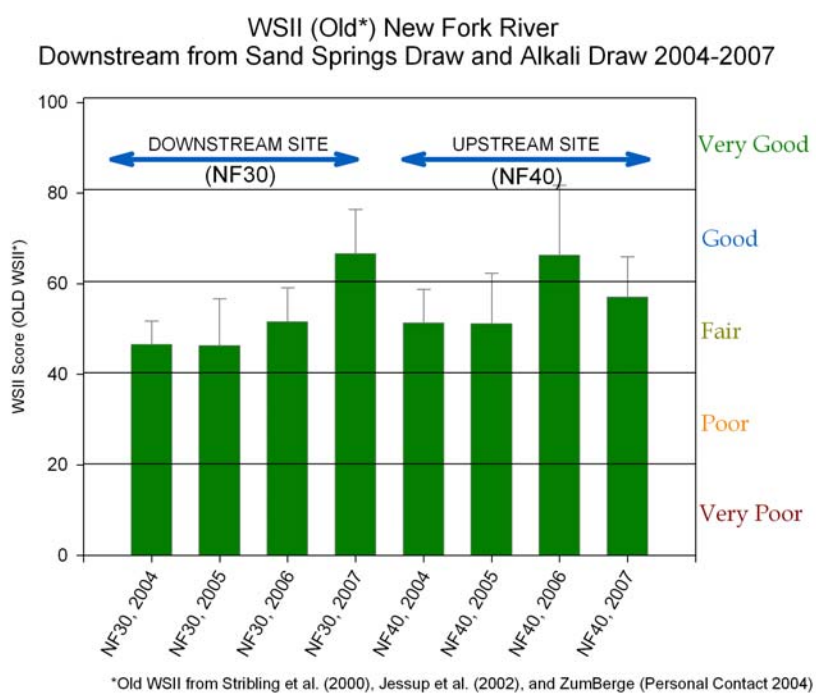


Figure 3.23. WSII-1. The composite samples indicated that there was no significant difference in the WSII-1 among years or between the sites NF30 and NF40.

Trends: Non-Insect Taxa

Aquatic insects are generally dominant in terms of richness and abundance in freshwater ecosystems of North America. When non-insects increase in abundance, it means there is something unusual occurring—usually sedimentation, intermittent flow, salinity or upwelling from subterranean sources. The abundance of non-insects in the samples has been the single most irksome metric throughout the entire study. It has suggested sedimentation problems in previous reports when there was no habitat to corroborate the finding. Additionally, the SS methods earlier in this report indicated that the non-insects were more abundant at NF30 than at NF17 and that they were correlated with the abundance of sand in the river. We needed to know if the finding for 2007 was due to new analytical methods, or if it occurred among the CS samples (upon which the historic data set is based).

Two-way ANOVA indicated that there was a significant difference over time (Years) and a marginally significant difference between sites NF30 and NF40 (Table 3.7) with the CS methods. NF30 continues to have abnormally high abundance of non-insect taxa—a pattern that started in 2004 (Fig 3.24).

Table 3.7. ANOVA % Non-Insects.

Dep Var: NonIn   N: 40   Multiple R: 0.531   Squared multiple R: 0.282

Analysis of Variance

Source	Sum-of-Squares	df	Mean-Square	F-ratio	P
SITE$	2874.726	1	2874.726	3.187	0.083
YEAR	7914.595	3	2638.198	2.925	0.048
SITE$*YEAR	925.723	3	308.574	0.342	0.795
Error	29764.638	33	901.959		

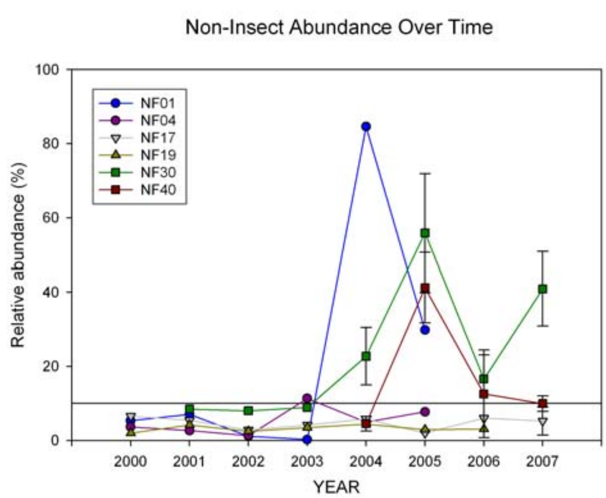


Figure. 3.24. Abundance of Non-insect Taxa in CS Samples. Site NF30 has consistently had abnormally high non-insect abundance since 2004.

4.0 Discussion

4.1. Influence of PAPA development on the New Fork River

Four separate analyses pointed to the differential response of NF30 to sedimentation. First, the analysis of habitat stratified by flow indicated that the sites sand content appeared to increase with water velocity rather than decrease. This usually indicates active sources of erosion that are not in equilibrium. Second, the amount of sand in the samples correlated with the abundance of non-insect taxa at NF30—even after correction for water velocity. The abundance of non-insects at this site was abnormally high in the SS samples. Third, the CCA ordination indicated that the overall taxonomic composition of the benthic community at NF30 (somewhat also at NF40) was highly influenced by sand content and embeddedness—and this included the abundance of aquatic worms—which should normally have very low abundance. Fourth, the ANOVA of sites over time, using the CS samples indicated that NF30 had significantly greater non-insect abundance than other sites, and that there has been an increase over time. The composition of the community structure was determined to be more sand-influenced than the samples collected from the sand-dominated East Fork River. Other metrics tested showed significant differences as well, but these were usually responding to same underlying changes—sediment inputs and abnormally high worm abundance at NF30.

The dominance of non-insects became apparent at several sites in 2004, and appeared to be returning to normal in 2006. However the combined 2007 results indicated that NF30 is under the influence of a unique, local, active erosion event. This should not incite panic by any means. Both versions of the WSII indicated that NF30 is in compliance with DEQ's healthy conditions (Figs. 3.22, 3.23) for Wyoming Basin streams; these data do not provide a foundation for regulatory intervention. Rather, your monitoring program has done what it was supposed to do: provide an early warning of changing conditions that might be related to development of the PAPA. Field reconnaissance should be able to identify the influential activities so that best-management practices can mediate further influence on the ecology of the New Fork River.

We did not study the source or nature of the sediment inputs to the river; particles finer than fine gravel, both sand and silt could occur, but only sand was definitively identified among the SS collections. The elevated relative abundance of collector-gathers at NF30, suggests that there are fine organic constituents to the sediment, but these are usually trapped among the grains of sandy substrata.

4.2. Next year's analyses

Next year the SS samples will be able to assess differences between 2007 and 2008 samples and in the context of the correlation with velocity and substrate composition. We will also develop an analysis of how to best standardize SS data so they are comparable with the CS sample data. This will allow the PAPA monitoring to phase out the CS, which are not as cost efficient as the SS samples. To phase out the CS without this analysis would devalue the historic data that have been amassed for this monitoring program.

The monitoring program has grown and adapted as needed. It could probably benefit from another subtle change in the site used as an upstream reference. Clearly the community of invertebrates at NF01 is an unattainable reference condition. This is not ideal, because it makes all the other sites appear to be in worse condition than they are and can obscure other important trends. This is especially important for the continued use of multivariate statistical methods. The study site schematic (Fig. 2.2) shows the approximate ideal location of the ideal upstream reference. This site would replace NF01, and should be selected especially to be representative of the condition of the river upstream from the PAPA, preferably downstream of Duck Creek and Willow Creek.

Literature Cited

Hargett, E.G. and J, R. ZumBerge. 2006. Devevelopment of the Wyoming Stream Integrity Index (WSII) for Assessing the biological condition of wadable streams in Wyoming. Wyoming Department of Environmental Quality, Cheyenne WY.

Hart D. D. and C. M. Finelli. 1999. Physical-biological coupling in streams: The pervasive effects of flow on benthic organisms. Annual Review of Ecology and Systematics. 30: 363-395.

Fonseca, D. M. and D.D. Hart. 1996. Density-dependent dispersal of black fly neonates is mediated by flow. OIKOS 75(1):49-58.

Marshall, B.D. 2005a. An Assessment of the Biological Condition of the New Fork River, near the Pinedale Anticline Project Area, Sublette County, Wyoming: 2004 Survey Results. Prepared for Sublette County Conservation District

Marshall, B.D. 2005b. An Assessment of Chemical and Biological Trends in the New Fork River Basin, Sublette County, WY: December 2005 Survey Results. Prepared for Sublette County Conservation District.

Marshall, B.D. 2006. An Assessment of the Biological Condition of the New Fork River, near the Pinedale Anticline Project Area, Sublette County, Wyoming: 2005 Survey Results. Prepared for Sublette County Conservation District and the Pinedale Anticline Water Task Group, Pinedale Wyoming.

Marshall, B.D. 2007. An Assessment of the Biological Condition of the New Fork River, near the Pinedale Anticline Project Area, Sublette County, Wyoming: 2006 Survey Results. Prepared for Sublette County Conservation District and the Pinedale Anticline Water Task Group, Pinedale Wyoming.

McCune, B., and M. J. Mefford. 1995. Multivariate analysis on the PC-ORD system. Version 2.0. MjM Software, Gleneden Beach, Oregon, USA.

Stribling, J.B., B. K. Jessup. J. Gerritsen. 2000. Development of biological and physical habitat criteria for Wyoming streams and their use in the TMDL process. Prepared for U.S. Environmental Protection Agency, Region VIII, Denver Co.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: a new eigenvector technique for multivariate direct gradient analysis. Ecology 67:1167-1179.

ter Braak. 1994. Canonical community ordination. Part I: Basic theory and linear methods. Ecoscience 1:127-140.

ZumBerge, J. 2004. Personal contact. Wyoming Department of Environmental Quality, Sheridan, WY. An MS Excel file of bioassessment criteria, modifying the WSII for all ecoregions.

Zar. J. H. 1999. Biostatisitcal Analysis (4th edition). Prentice Hall Pub. Upper Saddle River, NJ 07458.

Acknowledgements

The field work was conducted by SCCD, under the supervision of Kathy Raper. Laboratory work was conducted by certified macroinvertebrate taxonomists at EcoAnalysts /Inc., located in Moscow ID. The author would like to thank the board and staff of the SCCD for their editorial review of the report. The efforts of Daryl Walker and Kathy Raper were especially helpful. As usual, Sno Ann Engler's aid was essential to the smooth flow information among participants.